REPORT RESUMES

ED 012 066 24 CG 000 219

INTERPERSONAL ATTITUDES OF ADOLESCENTS. INTERIM REPORT.

BY- MOURSUND, JANET P. MATHIEU, PHILLIPPA L.

REPORT NUMBER BR-5-0496
REPORT NUMBER CRP-1621
EDRS PRICE MF-$0.27 HC-$5.96 149P.

DESCRIPTORS- HIGH SCHOOL STUDENTS, *GROUP COUNSELING, GROUP DYNAMICS, *SELF CONCEPT, COUNSELING EFFECTIVENESS, GROUP STRUCTURE, NONDIRECTIVE COUNSELING, *SOCIAL ATTITUDES, STATISTICAL DATA, RESEARCH, PERSONALITY ASSESSMENT, ATTITUDE TESTS, *INTERPERSONAL RELATIONSHIP, CALIFORNIA F SCALE, THE RELATIONSHIP INVENTORY, SOCIOMETRIC TEST, WAYS OF LOOKING AT PEOPLE, SELF REFERENT Q SORT, TEACHER RATINGS, MINNESOTA MULTIPHASIC PERSONALITY INVENTORY

THIS INTERIM REPORT PRESENTS DATA COLLECTED IN A STUDY OF THE INTERPERSONAL ATTITUDES OF ADOLESCENTS. TWELVE GROUPS OF HIGH SCHOOL JUNIORS FROM THREE DIFFERENT SCHOOLS MET IN COUNSELING SESSIONS. THE PRINCIPAL INDEPENDENT VARIABLES IN THE GROUPS WERE CONGENIALITY (STUDENTS CHOSE TO MEET WITH EACH OTHER) AND THE STRUCTURE LEVEL IMPOSED BY THE COUNSELOR. VARIOUS INSTRUMENTS WERE USED TO MEASURE BOTH THE PERSONAL AND THE INTERPERSONAL EFFECTS OF THE COUNSELING THROUGH PRE- AND POST-TREATMENT TESTING. RESULTS OF THE DATA ANALYSIS SHOWED CONFLICTS AND BORDERLINE SIGNIFICANCES WHICH SUGGESTED A NEED FOR FURTHER RESEARCH AND ANALYSIS. SOME CHANGES IN THE GROUPS DUE TO COUNSELING WERE APPARENT, BUT NONE COULD BE INTERPRETED AS DEMONSTRATING A CLEAR RELATIONSHIP BETWEEN VARIABLES OR STRONG SUPPORT FOR THE ORIGINAL HYPOTHESES. AT THIS TIME, DATA ANALYSIS IS INCOMPLETE AND A NUMBER OF PERIPHERAL INVESTIGATIONS BASED ON THE ORIGINAL DESIGN ARE CONTINUING. (NS)

ED012066

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

INTERPERSONAL ATTITUDES OF ADOLESCENTS

U.S.O.E. C.R.P. #1621

Interim Report

Janet P. Moursund, Ph.D.

Philippa L. Mathieu, Ph.D.

CG 000 219

Chapter I General Introduction and Review of Pertinent Literature
Chapter II Rationale and Instruments
Chapter III Plan of the Research
Chapter IV Tape Analysis Procedures
Chapter V Results and Discussion
Chapter VI Miscellaneous Additional Analyses
Chapter VII Conclusions

Chapter I

General Introduction and Review of Pertinent Literature

The discrepancy between the number of high school students capable of college work and the number actually entering college is of prime concern to educators and national administrators. Of no less concern are the large numbers of youngsters who enter college with no real notion of why they are there, only to drop out after a few disillusioning and wasteful months or years. It has been estimated that the number of potentially capable and contributing youths who are wasted with regard to college education may easily prove to be a major handicap in attaining the scientific and cultural status necessary for this country to continue in its role of world leadership.

Money alone is not the cause of failure to realize one's potential. We cannot say that those who drift into and out of college, neither setting nor realizing a goal, are the "spoiled rich"; many come from families in which real hardships are incurred that the child may have an education. Nor can lack of funds account for all of the failures to attend college where such attendance would be of value: the number of scholarships unclaimed yearly attest to this. It would seem that some more global attitudinal factor on the part of the potential student is at work. This kind of attitude set is of grave importance within the area of mental health as well as of education. Questions such as what the nature of these negative attitudes may be, how they are formed, and how they can best be replaced by more mature concepts, pose challenges to leaders in both fields.

It would appear that implicit in attitude formation is the effect on an adolescent of his peers, his family, his teachers, etc., in the formulation of his own impressions. It is as if his values at this time in his life are very much the result of his relationships with other people. He is much more vulnerable to impression by significant others than would be the case if he had had more experience: i. e., if he were an adult and not an adolescent. However, the commitment to act is one which he must carry out on his own, and the likelihood of retracing his steps at a later date becomes more remote with each passing year.

Through a better understanding of the kind of interpersonal relationships and perceptions held by the adolescent and the young adult, steps can be taken to insure a more complete realization of individual potential. Accurate assessment of interpersonal and social maturity can lead to more effective counseling techniques; to classroom experiences gauged to the student's level of comprehension and integration; to the provision of various "real life" situations which can enhance and encourage the individual's own efforts toward adulthood.

This study represents an exploration both of general attitudes and characteristics which have bearing on interpersonal relationships, and of specific interpersonal attitudes. The purpose of the research is threefold: to discover 1) how adolescents tend to perceive others in their world, 2) how these perceptions change over a limited period of time, and 3) how positive changes can best be facilitated in a general discussion framework. Knowledge of how attitudes such as these are formed, and with what degree of tenacity they are held, will facilitate both the traditional school guidance program and any new programs of social attitudinal exploration.

REVIEW OF THE LITERATURE

It is difficult, in reviewing even a limited portion of the research related to the present study, to know just where to begin. The literature on attitude change alone would fill several volumes. With this abundance of material in mind, we have limited ourselves throughout the following pages to consideration of only the most directly relevant studies.

According to Moscovici, writing in the Annual Review of Psychology (1963), it is not enough to consider the content of an attitude alone. Equally important is the framework within which it exists, and the "rejection or acceptance processes which are likely to organize and direct attitudes." One way of getting at these aspects is through a multiple test battery, which will measure a variety of personality and attitudinal factors at onec, and allow analyses of the interrelationships therein. Of particular interest in the context of the latter consideration mentioned by Moscovici, that of acceptance or rejection processes, is a study by Bass (1955) which suggests that the California F Scale is a valid measure of social acquiescence. Thus utilization of F Scale scores in conjunction with other psychometric and dynamic indicators of attitudes and attitude change should enhance our understanding of both the attitudes themselves and the likelihood that they will be acted upon.

The present research is primarily concerned with attitude change. Moscovici characterises this change as occurring through general "individual" processes (the quotes are his) or "through the agency of...specific psychological processes" which involve group action. This group action is the vehicle which has been chosen in the present investigation as the direction factor in attitude change, and the literature abounds in studies investigating the various aspects of the group which affect its functioning and outcome.

One of the most straight-forward appearing problems is that of the optimal size of the group. Bales and Borgatta (1955) studied groups ranging in size from two to seven members and found that as the size increased, the number of "tension release" statements and the number of "suggestions" offered also increased. Bass and Norton (1951) studied groups of two, four, six, eight and twelve subjects and found the greatest absolute variance in leadership assigned (e.g., in rank status) in the 6-group. Cartwright and Zander (1960) found that as size increases, so does heterogeneity, and with increased heterogeneity comes a lessening of cohesiveness. Castore (1962) reports that less than 9 members is optimal for the psychotherapy group, and Wolberg (1954) believes the optimal size for this sort of group to be from six to eight members.

The degree of experience in group interaction owned by the group leader is a significant factor in the overall functioning of the group. Mathieu and Moursund (1962), comparing a more experienced and a less experienced leader, found support for the hypothesis of greater positive change in groups meeting with the former. Yet here the question of confounding must be raised: to what degree did the orientation, techniques, etc. of these leaders also differ?

This leadership variable may be characterized as that of its "style." "Style" may be determined by the personal inclination of the group leader, by his training and professional orientation, by conditions imposed within the specific group situation, or by some combination of these factors. The now classic Lewin, Lippitt and White study (1939), which compared the effects of democratic, autocratic, and laissez-faire leaderships, set off a flood of similar investigations. In one of the more recent of these, Kipnis (1958) found differences in the degree of success experienced by a "participant" and a "directive" leader in their attempts to persuade children to change their reading habits. Miller and Biggs (1958) found relatively stable attitude changes as a function of undirected group discussion. On the other hand, Quay et.al. (1961) found that an authoritarian presentation of material to the group brought greater changes in attitudes than did discussion. It would seem desirable to attempt a resolution of at least part of this conflict of evidence. One technique which would greatly facilitate this resolution would be a more specific definition or delineation of the directiveness variable with regard to group leadership, inasmuch as "directiveness" can be interpreted to describe a wide variety of leadership behaviors.

Another aspect of the group which would seem relevant to its overall functioning is that of homogeneity of composition. Torrance (1961) found that groups which were homogeneous with regard to IQ and/or creativity showed more positive sorts of social interactions than did those which were heterogeneous. Mathieu and Moursund (1962) investigated tha changes in homogeneous and non-homogeneous groups (academic achievement as the criterion for homogeneity) during the course of time-limited counseling. Homogeneity appeared to create a climate leading to more commonly shared interpersonal perceptions; the effects of this variable with regard to other outcome criteria were somewhat masked by a confounding factor, that of absolute level of achievement within the group. Furthermore, examination of the raw data from this study indicates that the level of congeniality in the group also tended to interact with the above-mentioned variables.

Exline has investigated the congeniality variable. In a relatively non-confounded study, he found that the congenial group was superior to the non-congenial in terms of self-appraisal of group functioning (1957). This might indicate that the congenial group is better able to understand and evaluate the nature of the group interaction, and thus of their own role in that interaction. Within the Rogerian framwwork (Rogers, 1959) it is clear that this increased accuracy of perception is conducive to therapeutic gain. In partial support of such a contention, Lambert and Lowy (1958) have shown that discussion of issues within high-acquaintance groups reduces attitude variability, while little or no variability reduction occurs in low-acquainaance groups.

At this point the reader may well begin to question the general efficacy of the group in dealing with changes in attitude, personality, and the like. If there are so many interacting variables, he may ask, does not the group leader run the risk of undoing all he has accomplished by overlooking a small but important factor out of the welter of those he must consider? Would he not perhaps be better off to continue with the slower -- but better understood -- individual counseling interactions? In this conbext, it would be worth-while to consider for a moment the over-all efficacy of group techniques as they have stood up under a few representative experimental investigations.

Friedman (1960) used group meetings in connection with individual psychotherapy for high school students. She found that, though slow in getting started, the sessions were ultimately beneficial. Broedel _et.al_. (1960) report improvement in acceptance of self and others, and in interpersonal relationships, as a function of 16 sessions of group counseling with adolescents. Finally, Rhine (1960) reports enhancement of concept-attitude change by peer responces. These few studies cited, while only scratching the surface of relevant literature, do indicate that group work can definitely be an effective tool. It remains but to pin down the factors at work within the interaction.

Certain individual characteristics may also be relevant to the outcome of the group process -- that is, characteristics of the individual group member rather than of the group as a whole. The sex of the participating individual, for example, would seem relevant to the benefit he or she may obtain from the interaction. Tuma and Livson (1960) found girls more likely to accept authority (and thus to accept an authoritarian group structure) than boys. According to a study by Douvan (1960), adolescent girls were less concerned with values, more interested in interpersonal relationships, than were boys.

The adolescent's level of academic acheevement would seem to be a factor -- or at least a predictor -- of his value orientation. Thompson (1961) found high achievers to be oriented toward traditional values more than were low achievers. Mathieu and Moursund (1962) found low achievers to have significantly more authoritarian attitudes, as measured by the California F Scale, than high achievers. Thompson (1961) found that high ability students reacted less favorably to a counselor than did low ability students ("average" students had the most favorable reactions). It appears that, though achievement level is certainly a factor of importance in considering group functioning. The specific dynamics are yet to be delineated.

One further factor dealt with in some detail by the literature of group work is that of the interpersonal relationships and perceptions of the group members and their role, both as variables in the process and as an outcome criterion in and of themselves. Byrne (1961) found that strangers were rated higher in terms of intellect, morality, and adjustment if the rater thought the stranger had attitudes similar to his own. A study by Campbell _et.al_. (1961) showed that pre-adolescents' initial evaluation of their peers tended to shape their future actions toward the peers, rather than the actual behavior of the peers affecting later appraisals. O'Connor (1960) reports that disposition towards others, ability to evaluate others, and self-evaluation are all interrelated. Finally, Mathieu and Moursund found first impressions, whether good or bad, tended to be strengthened over the period of short-term group interaction. While all of these studies are suggestive, none of them has succeeded in isolating a specific dynamic of the interpersonal perception factor, nor in explicating the manner in which this factor affects the group outcome. It seems evident that further study is needed in the area.

In the present research, we have attempted to bring together a number of the variables discussed above, and to delineate the means whereby they interact in the ongoing counseling group, as well as their differential effects upon selected outcome criteria.

References

Bales, F. F., & Borgatta, E. F. Size of group as a factor in the interaction profile. In Hari, P., & Bales, R.F. (Eds.), *Small Groups*. New York: Alfred A. Knopf, 1955, pp. 396-413

Bass, B.M. Authoritarianism or acquiescence. *Journal of Abnormal and Social Psychology*, 1955, *51*, 616-623

Bass, B.M., Norton, Fay, & Tyler, M. Group size and leaderless discussions. *Journal of Applied Psychology*, 1951, *35*, 397-400

Broedel, M., Ohlsen, M., Proff, F., & Southard, C. The effects of group counseling on gifted underachieving adolescents. *Journal of Counseling Psychology*, 1960, *1*, 163-170

Byrne, D. Interpersonal attraction and attitude similarity. *Journal of Abnormal and Social Psychology*, 1961, *62*, 713-715

Campbell, J. D., & Yarrow, Marian R. Perceptual and behavioral correlates of social effectiveness. *Sociometry*, 1961, *24*, 1-20

Cartwright, D., & Zander, A. *Group Dynamics*. Evanston, Ill.: Row, Peterson & Co., 1960

Castore, G. P. Number of verbal interrelationships as a determinant of group size. *Journal of Abnormal and Social Psychology*, 1962, *64*, 456-488

Douvan, Elizabeth Sex differences in adolescent character processes. *Merrill-Palmer Quarterly*, 1960, *6*, 203-211

Exline, R. V. Group climate as a factor in the relevance and accuracy of social perception. *Journal of Abnormal and Social Psychology*, 1957, *55*, 382-388

Friedman, Alice Some notes on psychotherapy with a group of adolescents. *Acta Psychotherapeutica et Psychosomatica*, 1960, *8*, 147-151

Kipnis, D. The effects of leadership style and leadership power upon the inducement of an attitude change. *Journal of Abnormal and Social Psychology*, 1958, *57*, 173-180

Lambert, W.E., & Lowy, F. H. Effects of the presence and discussion of others on expressed attitudes. *Canadian Journal of Psychology*, 1958, *57*, 173-180

Lewin, K., Lippitt, R, & White, R.K. Patterns of aggressive behavior in experimentally created "social climates: *Journal of Social Psychology*, 1939, *10*, 271-279

Mathieu, Philippa, & Moursund, Janet *Relationship of group counseling to subsequent academic performance at the college level*. Cooperative Research Project No. OE 1417, U.S. Office of Education, Department of Health, Education, & Welfare. Madison, Wisconsin, 1962

Miller, K. M., & Biggs, J. B. Attitude change through undirected group discussion. *Journal of Educational Psychology*, 1958, 49, 224-228

Moscovici, Serge Attitudes and Opinions. *Annual Review of Psychology*, 1958, *49*, 224-228

O'Connor, W.F. The interrelationships of social perception, sociometric status, personality, and the ability to judge personality traits. USN Sch. Aviat. Med. Res. Rep., 1960, proj. #MR005.13-5001, Sub. 2, #9.iii

Quay, Lorene, Bartleet, C.J., Wrightsman, L.S. Jr., & Catron, D. Attitude change in attendant employees. *Journal of Social Psychology*, 1960, *51*, 27-31

Rhine, R.J. The effect of peer-group influence upon concept-attitude development and change. *Journal of Social Psychology*, 1960, *51*, 173-179

Rogers, C.R. A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A Study of a Science*, Vol. III *Formulations of the Person and Social Context*. New York: McGraw-Hill, 1959

Thompson, Muriel Students' evaluations of guidance services. *California Journal of Educational Research*. 1960, *11*, 167-169

Thompson, O.E. High school students' values: emergent or traditional. *California Journal of Educational Research*, 1961, *12*, 132-144

Torrance, E.P. Can groupings control social stress in creative activities? *Elementary School Journal*, 1961, *62*, 139-145

Tuma, E. & Livson, N. Family socioeconomic status and adolescent attitudes to authority. *Child Development*, 1960, *31*, 387-399

Wolberg, L.R. *The Technique of Psychotherapy* , New York: Grune & Stratton, 1954

Chapter II

Rationale and Instruments

Rationale

Although somewhat unorthodox in a report of this sort, a "personal history" of the project is perhaps the best way to communicate some of the ideas and processes which went into the logic and the structure of the study. Let us try, then, to trace through the kinds of questionings and conceptualizations which finally evolved into our research design.

In order to begin, we must go back past the "beginning" to a previous study. In the summer of 1962, the present authors completed a study of the effects of several kinds of group interactions on personal and social behaviors of college freshmen. The study (C.R.P. #1417) pointed to some interesting relationships; but, like so many others, seemed to open up as many new questions as it answered old ones.

Three types of group "interattions" were investigated in the 1962 study: group counseling, academically oriented group discussions, and a control type group which met only to study individually. Most intriguing were the data pertaining to the counseling groups. It seemed quite clear that something was happening to the students involved in these groups: reflected imperfectly, if at all, in the psychometric criteria of the research; shown more strikingly in the highly structured ratings made on the basis of tape recordings of the group sessions; and most apparent of all to the clinical intuition of the trained listener. It was our experience with this study -- the excitement of feeling that something was "there" and the frustration of being able to capture that something only in part in our experimental procedures -- that set us off on the present research.

What was needed was a more detailed and critical look at the group counseling interaction. The literature, discussed in Chapter I of this report, points to such a need; and the 1962 study further emphasized it. The question was, how best to proceed? The number of variables, both dependent and independent, which might be relevant is immense. Our final selection from among this array of variables was based in part upon the findings and suggestions to be found in the literature, in part upon various studies in progress at the Wisconsin Psychiatric Institute which suggested profitable avenues of study, and finally upon our own particular curiosities about certain areas.

The age group with which it was finally decided to work was that of high school juniors. Several factors dictated this choice. High school students seem to be an inherently interesting group, inasmuch as they constitute one of the largest potential populations in need of guidance services. They are also volatile, receptive to new ideas, and seem to have the virtue of being relatively unsophisticated with regard to psychological research (in contradistinction to the ubiquitous "elementary psychology student" so often used as a research subject) while at the same time intellectually capable of assimilating the kinds of concepts important to positive personal growth. Eleventh-graders were chosen because they were not new to the high school routine, as many 10th-graders would be; but yet would be in the school system and available for follow-up studies during the following year.

Volunteer subjects were used throughout the study. This was partly a matter of expediency. However, the clinical data of the previous study had indicated the much greater likelihood of positive change occurring within a time-limited framework on the part of non-resistant subjects, and it was felt that using volunteers would facilitate the observation of those variables in which we were primarily interested.

The two main independent variables which were chosen for sutdy were structural in nature: the internal structure of the group, and the external structure imposed upon the group by the group leader or counselor. The internal structure variable was characterized as "congeniality" among the members. Some of the groups were organized as "congenial": the members knew and liked each other and chose to work together in their group. The other groups were "non-congenial" (not to be confused with "uncongenial"); these were composed of students who had not explicitly requested to work with each other, and, in some cases, were not even acquainted prior to bhe beginning of the group sessions.

The externally imposed structure was rather rigorously specified to the group counselors as a part of their initial instructions. Three structure levels were imposed. The highly structured groups were given discussion subjects for each meeting, and the group leaders were supplied with specific lists of questions related to those subjects to which the group was supposed to address itself. The moderately structured groups had general topics for each meeting, but beyond that were free to proceed as they wished. The low structure groups were allowed to pursue any avenue of discussion they wanted, within the discretion of the group leader; it was assumed that these groups would follow the general pattern of "non-directive" counseling.

At this point in our thinking, we had conceptualized a number of groups of adolescents, meeting in small groups under various structural conditions, but all undergoing some sort of experience which could be classified generally as "counseling." We fully expected that the experience would be a positive one for these young people, and that there would be enhancement of various aspects of their behaviors as a result of the experience. The next problem to be attacked, then, was that of measuring these changes in some concrete, communicable, and manipulable fashion. Again, we had a multitude of possible variables to consider.

It was finally decided to divide the kinds of effects to be looked for into two main categories: personal and interpersonal. Our emphasis throughout has been upon interpersonal functioning; but it is recognized that this interpersonal behavior grows out of the individual personalities who are interacting. Effects of the group experience on the group members as individuals, therefore, could not be ignored. The primary concept upon which we focussed here was that of the self-image. Self-image as a central factor in personality and personality change has been emphasized by many theoriests (Rogers, Snygg & Coombs, Brookover, G. H. Mead); we looked closely at descriptions of both the "real" and the "ideal" self, and their relationships. A measure of authoritarianism, or, more generally, rigidity, was also considered to be an intra-personal variable which would have great bearing on a subject's interpersonal behavior.

Various instruments were used to explore facets of interpersonal functioning. These ranged from teacher ratings through relatively well-documented and structured instruments to scales devised by the authors specifically for the study, and will be described in detail in the following section.

One of the most central aspects of the study has been the tape-recording of the group sessions themselves. These have been segmented and analyzed by means of techniques developed at the Psychiatric Institute. They thus provide not only a statistically useful body of data, but a rich source of clinical information as well.

Minnesota Multiphasic Personality Inventory

Inasmuch as one of the major areas of possible positive change which are to be examined in this research is that of personality, it would seem necessary to utilize an instrument for the measurement of personality. "Personality tests" of one sort or another exist today by the score, but they may be classified roughly into two sorts: the structured and the nonstructured (projective). Due to such considerations as length of time needed for administration, difficulty in scoring, poor reliability and/or validity data, and questionable theoretical foundations, most if not all of the latter variety can be ruled out as satisfactory tools for an investigation of the present sort.

Among the structured tests of personality, the Minnesota Multiphasic Personality Inventory has probably received more attention than any other instrument (Rev. Ed. Res. 1959 p. 62). Ellis (1946) suggested that this test alone (of all individually administered personality scales) might be valid. Both the great popularity of the MMPI, and the availability of validity and reliability data on it and its sub-scales, recommended it to our use.

The MMPI was constructed by means of empirical item selection (Welsh & Dahlstron, 1956) and therefore it is difficult to assign a theoretical basis by means of which it may be assessed. Also, since it has been validated and utilized primarily with psychotic/neurotic populations, the validity of its application to "normal" S's may be called into question. One possible solution to this problem lies in the technique of factor analysis: if the test, or some sub-scale of the test, can be shown to depend heavily on certain factors, this provides the beginning of a theoretical framework and an interpretive "handle" for the researcher or clinician. One such study has been carried out by Comrey (1957), working only with the items on the D (depression) scale of the MMPI. The main factor on this scale, according to his findings, is that of "neuroticism"; he was also able to extract and label such factors as cynicism, religious fervor, poor physical health, hostility, depression. These factor labels, while only descriptive, provide a relatively detailed picture of the individual who tends to score high on the D scale.

Another study of the D scale was carried out by McCall (1958). His primary interest was in the item validity of the scale, and he found items to be valid in proportion to their face validity. While this raises rather interesting questions as to "fakeability," it nevertheless would seem to indicate that, for McCall's subjects at least, the obvious clinical flavor of the items did not interfere with their usefulness.

There is much to be desired in the MMPI as a research measure of personality change. Types of changes to be expected as a concommitant of therapeutic progress are not clear; indeed, the criteria by which "therapeutic progress" shall be determined are often ambiguous. Nevertheless, the MMPI has much to recommend it: objective construction, conscientious validation and reliability studies, and a wealth of experimental data related to its possible applications. While any conclusions as to an individual positive personality factors based solely on the MMPI data must of necessity be tentative ones, the instrument can provide a basis for hypothesizing such factors and, in the presence of other corroborative data, contribute significantly to the assessment of total personality movement.

The MMPI has been used in the present study as an exploratory and hypothesis-generating tool rather than as a means of hypothesis testing. It was administered only once, prior to the experimental treatment, and to only part of the sample. It was felt by the administrative staff of the parochial high school that the instrument in its entirity was not appropriate for administration to members of their student body. Therefore, the test was taken by only the students from the public and the university high schools.

During the testing of these subjects, it was found that the MMPI in particular elicited an extreme amount of resistance and/or anxiety. One subject, in fact, was unable to complete the test because he felt that the questions were designed almost with him in mind, and he was too upset to continue. Both the abnormal-clinical nature of many of the questions and the length of the instrument (600+ questions) disturbed the students.

Ways of Looking at People

The WLP (Ways of Looking at People) Scale is one of the instruments which were designed primarily for use in the present research. Although the WLP has a rather complex history, it was originally conceptualized as a means of measuring social attitudes and/or social maturity in adolescents. It has kept this flavor through its several revisions.

"Attitudes," of course, are as difficult to measure as they are to define. Rather than become involved in the complexities of the various theoretical arguments "for" and "against" the validity of the many attitude schemata which have been suggested, we decided to take a more pragmatic point of view. We therefore attempted to construct a scale which would concern itself with those attitudes and values which we felt to be directly related to the kinds of social maturational processes focussed on in this research.

The first step in constructing the scale was a simple listing of areas of thought in which one might expect social values and attitudes to form. These included such areas as loyalty to one's friends; pity for others as opposed to blame and punishment for others; perceived similarity to adults, and self-confidence. When this list had been pared down and collapsed into about eight areas, specific questions were written within each area. These were questions which could be answered according to a 5-point Likkert type scale: strongly agree, agree, don't know, disagree, or strongly disagree. A few representative questions follow.

> Do you think people basically like each other?
> Are you pretty much like everyone you know?
> Do you need people around you to be happy?
> Can people really be honest with each other?

The questions were so constructed as tobe thought-provoking, and to have no obvious "right" or "wrong" answer.

The first pilot study for the WLP consisted of administering it to a group of church young people. The results indicated quite clearly a basic flaw in the scale. We had succeeded so well in constructing "provocative" questions with many possible interpretations that there was absolutely no consensus whatsoever in the pilot group. The young people in the pilot group, who might have been expected to have somewhat similar frames of reference, showed no discernable patterns in their test responses. Yet the ideas and concepts seemed to intrigue the students, who expressed great interest in the scale itself. We were faced with the necessity of revising the scale in such a way that it would have the potential to reflect some consensus or convergence of values as the subject became more socially mature; yet would still retain the flavor which made it interesting to the respondee.

The first revision shuffled and added items and finally emerged with ten categories which were perhaps less overlapping than in the original. As the result of a conscious shift in our modus operandi, the revised questions were much less ambiguous; that is, the "right" (or at least the "do-gooder") answer was fairly evident on many items. Some examples:

> Adults seem to feel different than I do about most things.
> People who do bad things should always be punished.
> I sometimes think I'm just not as good as most other people.
> People are just out to get what they can from you.

These items, in addition to being somewhat simpler than in the original version, were also phrased as statements rather than questions. In order to avoid response-set effects, some of the items in each category were worded so that the predicted "more mature" answer would be in the negative rather than the positive direction. These negative item answers were reversed in the scoring procedure.

The revised WLP was administered to a normative sample of about 2500 students. These students were currently enrolled in a public senior high school in Denver, Colorado; a private boys' secondary school in Philadelphia, Pennsylvania; and a Catholic high school in Boston, Massachusetts. The normative data were split randomly into two parts, balanced for sex, age, year in school, and geographic area. From the first half of the data, final selection of "score" test items and categories was made. The second half was used as a validation sample for this final item selection.

From the first half normative data, item response tallies were made on the basis of age, grade in school, sex, and post-high school plans. In other words, the normative group was first divided into age levels: 14-, 15-, 16-, 17-, and 18-year-olds. For each WLP item a tally chart was made showing the number of students in each of the age groups who responded --, -, 0, +, and ++. This was repeated with the sample divided into 9th-, 10th-, 11th-, and 12th-graders; into boys and girls; and into those planning to attend college, planning to continue their education but not in college (secretarial school, barber school, etc.), those planning to take jobs immediately after graduating, and those with no specific post-high school plans. Items were selected or discarded on the basis of discriminating among these various sets of sub-groups.

It was found that age and grade in school provided by far the most consistent criteria for item selection. Items which satisfied these criteria relatively well (choice being, at this point, on the basis of inspection rather than any more rigorous statistical procedure) were arranged in clusters according to the value category to which they were assigned, and categories which did not have a number of discriminating items were discarded.

Within the non-discarded categories, inter-item correlations were run on all items (both discriminating and non-discriminating) (see Figure II-B-1). The degree of correlation within the category matrix served as another selection criterion. Thus, an item which did not seem to discriminate well on the normative sample, but which was strongly related to several other items within its category which did discriminate, was selected as a "score" item.

The final version of the scale looked on the surface exactly like the first revision; the non-score items were left in as filler items. As can be seen from the protocol in Appendix B, items belonging to a given category were not together on the scale, but were scattered throughout the whole list.

Eight "score" categories remain in the final revision of the WLP. "Similarity to Adults" contains relatively straightforward items relating to the degree to which the respondee feels himself to be "like" an adult. "Giving and Taking" deals with a variable which might be characterized as "selfishness-unselfishness": helping others at the expense of one's self, responsibility to others, etc. "Pity and Blame" is concerned with a permissive, forgiving attitude as opposed to a strict eye-for-an-eye philosophy. "Basic Values" is perhaps misleadingly titled; it deals not with the nature of the respondee's values, but rather with the importance which he places upon awareness of these values. Another relatively straight-forward category is "Confidence," which is concerned with the respondee's estimation of himself and his capabilities. "Liking Others" deals with general sociability and friendliness. "Trust and Mistrust" elicits information about the respondee's opinion of others' trustworthiness. Finally, "Basic Nature of People" asks whether people in general are "good" or "bad."

In most cases, the normative sample confirmed our expectations as to the direction of the "mature" answer. However, this was not always true. Where the data belied the theoretical expectation, the final direction of the scoring was determined by the actual responses of the normative group.

Once the final scoring procedure had been determined, the second half of the normative data was used as a validation group. Differences between age groups (the year-in-school grouping correlated so highly with age that it was dropped as a validation criterion), sexes, socioeconomic levels, and regional groupings were tested by means of t tests and analyses of variance.

In order to obtain a large n for the analyses, the first age group analysis had only three levels: 15-, 16-, and 17-year-olds. On four of the eight categories there were significant score increases with age. Three of the other categories showed this trend at a non-significant level; there was a non-significant trend in the opposite direction in one category, "Giving and Taking" (see Figures I-B-2 through I-B-10).

Dropping the within-group n to 13 allowed testing the full age range of 12-year-olds through 19-year-olds. When this was done, the significant differences in each category disappeared. This seems to have been due mainly to excessive within-group variance, probably caused at least in part by sex differences (see below). However, even with these smaller groups, there was some evidence that the trend of score increase with age holds for most of the categories. One notable exception was the 12-13 year group, which appeared consistently higher than would be expected theoretically.

With regard to sex differences, girls scored significantly higher than boys on six of the eight categories. On one of the remaining two, girls were higher but not significantly so. Boys scored higher on only one category, "Confidence"; and the difference was not significant (see Figure I-B-11).

Two of the categories yielded a significant sex by age interaction. One of these, "Pity and Blame," shows no consistent relationship and is therefore difficult to interpret. The other, "Liking Others," shows girls consistently high at ages 15, 16, and 17; while boys start lower and approach, at 17, the level held by the girls. This may reflect the earlier socialization and group dependence often seen among girls (see Figure II-B-12).

There was a clear trend for the middle category of socioeconomic status rankings to score higher than either the high SES group or the low SES group. The differences were significant on three categories, and on only one category "Confidence," was the trend not upheld (see Figures II-B-12-21). It is interesting to note that the "Confidence" category was also the only one not to follow the boy-girl difference trend. With regard to the SES differences, we have theorized that the items may be written in such a way as to place high valuation on traditional "middle-class" values; this would create the SES differences which were found to exist.

The differences between schools in WLP scores were quite clear and consistent for most categories. In 6 of the 8 categories, the students from the parochial school in Boston scored highest, and the score differences were significant; in another the trend held but the difference was not statistically significant (see Figures II-B-22-23). Only one category, "Confidence," did not place the parochial students high; here they scored significantly lower than the other groups. Again, it should be noted that the "Confidence" category has consistently showed up at variance with the other 7 categories of the WLP. Of the six categories in which parochial students scored significantly higher, four showed the Philadelphia private school students to score lowest. In category "Confidence" the private school students were significantly higher than the others.

It is unfortunate that the confounding of the regional variable is such as to obscure any interpretations of these results. Observed differences may be related to geographical area, to type of school, or both. It seems reasonable to hypothesize, at this point, that the emphasis on moral and social values often found in a parochial school was an important factor in the high scores of the Boston area subjects.

Thus the WLP comes to the present research with a good deal of supportive evidence as to its validity. While this validity (as a measure of that construct which we call "social maturity") cannot be said to have been proven, the instrument seems at least to address itself to the general area with which we are concerned; and at best to tap a variable or set of variables which have hitherto remained unmeasured.

Category: Similarity to Adults

	21	34	43	46
15	.196	.185	.527	.186
46	.087	.059	.248	
43	.199	.317		
34	.001			

Category: Giving and Taking

	10	39	31	27
12	.087	.113	.320	.187
27	.158	.127	.184	
31	.115	.200		
39	.195			

Category: Liking for Others

	47	19	28	14
7	.110	.268	.334	.266
14	.056	.235	.138	
28	.215	.195		
19	.114			

Category: Trust and Mistrust

	50	48	35	39	5
3	.208	.100	.296	.353	.287
5	.200	.082	.132	.230	
30	.135	.320	.272		
35	.225	.256			
43	.256				

Category: Pity and Blame

	4	40	16	22
11	.222	.113	.197	.066
22	.228	.236	.134	
16	.187	.091		
40	.072			

Category: Confidence

	42	36	24	17
2	.232	.294	.164	.167
17	.046	.119	.203	
24	.065	.165		
36	.227			

Category: Basic Values

	44	53	33	20
6	.105	.132	.240	.128
20	.114	.393	.232	
33	.383	.336		
53	.166			

Cateogry: Basic Nature of People

	51	37	23	32
8	.114	.355	.164	.156
32	.318	.274	.366	
23	.348	.280		
37	.234			

Figure II-B-1. Item intercorrelations within 8 categories of the WLP. (marginal numbers refer to item numbers)

	SS	df	MS	F	p
Age	199.64	2	99.82	7.27	<.01
Sex	37.99	1	37.99	2.77	<.10
Age x Sex	16.95	2	8.47		
Error	3460.23	252	13.73		

Figure II-B 2. Analysis of variance for category "Similarity to Adults" of the WLP.

	SS	df	MS	F	p
Age	75.29	2	37.64	2.87	<.10
Sex	126.98	1	126.98	9.69	<.01
Age x Sex	74.91	2	37.46	2.86	<.10
Error	3301.35	252	13.10		

Figure II-B-3. Analysis of variance for category "Pity and Blame" of the WLP.

	SS	df	MS	F	p
Age	14.05	2	7.02		
Sex	101.72	1	101.72	9.54	<.01
Age x Sex	16.28	2	8.14		
Error	2685.86	252	10.66		

Figure II-B-4. Analysis of variance for category "Giving and Taking" of the WLP.

	SS	df	MS	F	p
Age	24.01	2	12.00	1.13	
Sex	25.43	1	25.43	2.39	
Age x Sex	31.42	2	15.71	1.48	
Error	2680.61	252	10.64		

Figure II-B-5. Analysis of variance for category "Confidence" of the WLP.

	SS	df	MS	F	p
Age	37.50	2	18.74	1.99	
Sex	197.97	1	197.95	21.03	<.01
Age x Sex	66.10	2	33.05	3.51	<.05
Error	2372.51	252	9.41		

Figure II-B-6. Analysis of variance for category "Liking Others" of the WLP.

	SS	df	MS	F	p
Age	31.54	2	15.77	1.34	
Sex	212.23	1	212.23	18.09	<.01
Age x Sex	24.35	2	12.17	1.04	
Error	2956.84	252	11.73		

Figure II-B-7. Analysis of variance for category "Basic Values" of the WLP.

	SS	df	MS	F	p
Age	82.15	2	41.07	3.03	<.05
Sex	236.47	1	236.47	17.45	
Age x Sex	16.66	2	8.33		
Error	3415.72	252	13.55		

Figure II-B-8. Analysis of variance for category "Trust & Mistrust" of the WLP.

	SS	df	MS	F	p
Age	122.80	2	61.40	45.62	<.01
Sex	113.34	1	113.34	8.42	<.01
Age x Sex	42.02	2	21.01	1.56	
Error	3392.28	252	13.46		

Figure II-B-9. Analysis of variance for category "Basic Nature of People" of the WLP.

	Age		
Scale	15	16	17
S-A	16.33	15.85	17.91
P/B	15.20	16.27	16.41
G/T	19.05	18.51	18.61
C	16.88	17.63	17.31
L-O	18.01	18.00	18.80
B-V	19.38	19.71	20.23
T/M	17.67	17.54	18.79
B-N	18.42	17.67	19.36

Figure II-B-10. Mean scores for 3 age groups on 8 categories of the WLP.

	Sex	
Scale	Male	Female
S-A	16.31	17.08
P/B	15.26	16.66
G/T	18.09	19.35
C	17.59	16.96
L-O	17.39	19.14
B-V	18.87	20.68
T/M	17.04	18.95
B-N	17.82	19.15

Figure II-B-11. Mean scores for Males and Females on 8 categories of the WLP.

	Age			
Sex	15	16	17	
Male	15.21	14.98	15.58	Category "Pity and Blame"
Female	15.19	17.56	17.23	

	Age			
Sex	15	16	17	
Male	16.56	17.02	18.58	Category "Liking Others"
Female	19 47	18.93	19.02	

Figure II-B-12. Means for Age x Sex interaction on two categories of the WLP.

Source	SS	df	MS	F	p
Age	94.39	5	18.88	1.35	
SES	12.11	2	6.06		
Age x SES	86.25	10	8.63		
Error	3023.23	216	14.00		

Figure II-B-13. Analysis of variance for category "Similarity to Adults" the WLP - Age x SES.

Source	SS	df	MS	F	p
Age	81.41	5	16.28	1.36	
SES	78.47	2	39.24	3.28	<.05
Age x SES	241.99	10	24.20	2.03	<.05
Error	2580.61	216	11.95		

Figure II-B-14. Analysis of variance for category "Pity and Blame" of the WLP.

Source	SS	df	MS	F	p
Age	60.18	5	12.04	1.02	
SES	83.60	2	41.80	3.54	<.05
Age x SES	253.02	10	25.30	2.14	<.05
Error	2551.69	216	11.81		

Figure II-B-15. Analysis of variance for category "Giving and Taking" of the WLP.

Source	SS	df	MS	F	p
Age	44.55	5	8.91		
SES	16.06	2	8.03		
Age x SES	52.76	10	5.28		
Error	2299.23	216	10.64		

Figure II-B-16. Analysis of variance for category "Confidence" of the WLP.

Source	SS	df	MS	F	p
Age	51.10	5	10.22		
SES	51.10	2	25.54	2.12	
Age x SES	122.50	10	12.25	1.01	
Error	2606.77	216	12.07		

Figure II-B-17. Analysis of variance for category "Liking for Others" of the WLP.

Source	SS	df	MS	F	p
Age	61.88	5	12.38		
SES	107.70	2	53.85	4.08	<.05
Age x SES	143.17	10	14.32	1.09	
Error	2849.69	216	13.19		

Figure II-B-18. Analysis of variance for category "Basic Values" of the WLP.

Source	SS	df	MS	F	p
Age	41.93	5	8.39		
SES	43.18	2	21.59	1.39	
Age x SES	235.23	10	23.52	1.52	
Error	3351.54	216	15.51		

Figure II-B-19. Analysis of variance for category "Trust and Mistrust of the WLP.

Source	SS	df	MS	F	p
Age	55.36	5	11.07		
SES	119.86	2	59.93	4.04	<.05
Age x SES	190.66	10	19.07	1.29	
Error	3204.46	216	14.84		

Figure II-B-20. Analysis of variance for category "Basic Nature of People of the WLP.

Category	SES Level		
S-A	16.89	17.31	16.78
P/B	15.60	16.77	15.49
G/T	17.94	19.33	18.26
C	17.63	17.33	16.99
L-O	17.41	18.54	17.81
B-V	18.27	19.80	19.60
T/M	17.27	18.32	17.76
B-N	17.60	19.31	18.10

Figure II-B-21. Socioeconomic status level means for 8 categories of the WLP.

Scale		SS	df	MS	F	p
S-A	Area	32.58	2	16.29	1.22	
	Error	3173.68	237	13.39		
P/B	Area	67.06	2	33.53	2.83	<.10
	Error	2811.68	237	11.86		
G/T	Area	247.11	2	123.55	9.56	<.01
	Error	3063.39	237	12.92		
C	Area	143.66	2	71.83	5.56	<.01
	Error	2588.28	237	10.92		
L-O	Area	207.56	2	103.78	8.93	<.01
	Error	2751.44	237	11.61		
B-V	Area	314.43	2	157.22	12.22	<.01
	Error	3050.30	237	12.87		
T/M	Area	347.66	2	173.83	13.75	<.01
	Error	2995.14	237	12.64		
B-N	Area	345.41	2	172.70	13.44	<.01
	Error	3045.09	237	12.85		

Figure II-B-22. Analyses of variance for 3 geographical areas on 8 scales of the WLP.

	Area		
Category	Boston	Denver	Philadelphia
S-A	17.46	16.61	17.30
P/B	16.89	16.36	15.60
G/T	19.73	18.86	17.28
C	16.49	16.80	18.26
L-O	19.38	18.14	17.10
B-V	20.83	19.30	18.03
T/M	20.10	17.31	17.88
B-N	20.65	18.06	18.15

Figure II-B-23. Geographic area mean scores on 8 categories of the WLP.

California F Scale

In a study of intrapersonal and interpersonal behavior, the concept of "attitude" must invariably be considered. Attitudes, however they may be defined, influence nearly every sphere of individual and group behavior. An individual's personality both grows out of and contributes to his attitudes about himself and others. One's attitudes toward people in general and toward specific individuals in particular play a vital role in determining one's social behavior. And, finally, "attitude" as a theoretical construct is so interwoven with the other aspects of "personality" (as another theoretical construct) that only by means of verbal and definitional restrictions can they be separated at all.

The concept of authoritarian rigidity seems to be perhaps most clearly related to the hypotheses of change implicit in this study, and the California F Scale was the instrument used as a measure of this attitude.

The large amount of research with the F Scale, reported in the literature, provides ample basis for evaluation of the scale's reliability and validity. The original description of the scale's constructuon (Adorno, 1950) describes in detail the means by which the items were validated; Lambert (1958) has commented on the care which went into this procedure. Zuckerman and Oltean (1959) report a significant correlation between the F Scale and the Authoritarian-Control factor of the Parent Attitude Research Instrument (PARI); similarly, Hart (1957) found a correlation of .63 between F Scale scores and self-reported use of "non-love" oriented material discipline techniques. Also related to this aspect of authoritarian behavior are the findings of Sheldon, Coale and Copple (1959), who report that teachers judged to be "warm and friendly" score significantly lower on the F Scale than do those not so rated.

More interesting than evidence that the F Scale is a measure of authoritarianism *per se*, however, are the studies relating F scores to a constellation of variables including perceptual and conceptual rigidity, stimulus bondedness, ego defensiveness, etc. Jackson (1959) describes the stimulus-bound quality of high F score subjects in terms of their inability to keep a Nekker cube from reversing or "flipping." The utilization of information in making accurate judgments was studied by Lipitz (1960), who found that low F subjects were considerably more flexible in incorporating information into their judgments than were high F subjects. Cohn (1957), while questioning the value of the F Scale as a definite measure of prejudice and/or facism, cites a number of studies in which the scale does relate significantly to attitudinal rigidity. McClintock (1958) used the F Scale to discriminate groups of subjects who reacted to "change inductions" in the hypothesized directions: in other words, he was able to predict his subjects' reactions to several kinds of communications on the basis of their F scores. Finally, Mozar (1960) reports that high F subjects tend to score high on the semantic differential test; e.g., to take an extreme position.

The patterns of reactions described in the foregoing paragraph seem to be closely related to the descriptions of the early stages of therapeutic process in Rogers' theory of personality change (Rogers, 1961). According to this theory, such patterns of rigid, closed, stimulus-bound and well-defended behavior should change if process occurs. If any of the group interaction procedures are beneficial, in a therapeutic sense, this benefit should be reflected in a change in F scores. Thus not only in the area of attitudes *per se*, but also in the personality-determined patterns of response-to-world, the F Scale would seem well-suited to reflect change and growth.

Reliability figures on the F scale in its several revisions indicate that the measure is a relatively stable one. Adorno (1950) reports increased reliability on each of his revisions, with the reliability quotient of the final version equal to .90.

Perhaps the most serious criticism of the F scale is the problem of response set. Because all of the scored items are phrased in the same direction, it has been suggested that a bias to answer in one direction or the other might influence the subject's score significantly without being actually a reflection of the attitude in question. Adorno offers three comments on this criticism: 1) the low inter-item correlations found in validation studies indicate that response bias does not act significantly on the test as a whole; 2) from the data which was gathered, it was felt that subjects actually tried to vary their answers so as not to appear "extreme" in their views; and 3) similar scores were obtained on scales which used both positive and negative items. It has also been suggested that an acquiescence set, rather than tending to invalidate the scale, may actually contribute to its validity inasmuch as acquiescence may itself be an important factor in the general cluster of authoritarian attitudes.

Self Referent Q Sort

Q technique was developed by Stephenson (1953) in an effort to make methodologically possible the study of the single case. It has subsequently become much more widely applicable, being used to speak to the issues of "the objectivity of 'subjectivity,' the claims of nomothetic versus idiographic principles, the definition of samples, proper regard of psychological types" and many more.

As it has been used in this study, the Q sort consists of 80 statements about the self, ranging from extremely negative (e.g., "I am a failure") to extremely positive (e.g., "I am a strong competent person"). Subjects sort these statements, printed on individual cards, into a nine-place forced-normal distribution. From this distribution, a score based on an "ideal-expert" sort is derived, which is taken as a measure of adjustment (Rogers & Dymond, 1953; Lewis, 1959).

Butler and Haigh, in their exposition of the use of the Q technique as a means of assessing the success of psychotherapy (1954), have pointed out the interrelationship of sets of single self-perceptions existing for an individual. It would seem that allowing a subject to rank-order the validity of self-referent statements would be a more valid indicator of such an interrelationship than simply asking the subject to rate them as more or less true. Thus the sorting technique itself offers more latitude in the expression of the subjects' self-perceptions than does a questionnaire type instrument.

One question which must be raised with regard to any measure of psychological variables is that of reliability. Is the Q sort a reasonably reliable instrument? And, concommitantly, is the "self-concept" sufficiently stable in itself to permit its treatment as an ongoing and measurable aspect of personality? Two recent studies support the reliability of the Q sort. One, by Frank (1956), reports a test-retest reliability of .93 to .97 on a group of 10 subjects. The other, by Engel (1959), used an adolescent population; the author concludes that the self-concept (as measured by Q techniques) is "relatively stable" over a two-year test-retest period.

A further criticism which has been made of the Q-sort is that the forcing of the sort into a normal distribution may adversely effect the validity of the score. A study was carried out by Block (1956) in an effort to clarify this point; he reports that over a series of comparisons the forced sort was either "equal or superior to" a similar but unforced sort method.

Finally, it is reasonable to ask whether the Q-sort adjustment score can, in fact, reflect the kind of change which might be expected to occur as a function of successful therapeutic interaction. Dymond (1954), reporting on the extensive University of Chicago Counseling Center research carried out under the direction of Carl Rogers, describes the change in the experimental (therapy) subjects observed in that study: "the mean adjustment score of the total experimental group after therapy was significantly higher than their pre-therapy score."

Thus, from the literature, it would seem reasonable to conclude that the Q sort can be a valid measure of change, and that it provides a reliable reflection of the subject's self-perception.

The Relationship Inventory

G. Barrett-Lennard, working at the Counseling Center of the University of Chicago, became interested in 1960 in developing a scale which would measure the client's perceptions of those aspects of therapist behavior which Rogers (1957) had postulated to be "necessary" and "sufficient" to facilitate therapeutic change in a client. The Relationship Inventory is the result of Barrett-Lennard's work (Barrett-Lennard, 1962).

The Relationship Inventory, in its present form, consists of 72 items and four separate scales: Positive Regard, Empathic Understanding, Congruence, and Unconditionality of Regard. Each of these scales is designed to measure the perception of one aspect of interpersonal behavior which, theoretically, should lead to positive personality change as an outgrowth of the behavioral interaction. Although the instrument was specifically developed for use in a formal therapy setting, Rogers' later formulation justifies its use as a measure of these characteristics in other interpersonal situations.

That the Inventory is indeed a valid discriminator between personally meaningful and nonmeaningful relationships is supported by a study by Berlin and Gendlin (1960). Using as subjects a group of sorority girls, the authors measured both the perception of the Conditions and the physiological concommitants of a verbal interaction (polygraph recordings) between subjects who had a "good and close" relationship and who had a "cold and distant" relationship. Both the Relationship Inventory and the polygraph data showed distinct differences between the two kinds of relationships; moreover, the relationship between data from the two instruments was clearly indicated.

A more recent study (Van der Veen, 1962) has demonstrated that it may be the client's perception of the Conditions, rather than their actual presence in the relationship (as judged by observers and by the therapist himself), which tends to facilitate therapeutic change. In other words, it would seem that the person on the "receiving end" of the Conditions must be at least minimally aware that they are present in order to benefit from them. Thus, if an individual's perception of the presence of the Conditions in a given relationship can be measured accurately, this should provide an index of the degree to which that relationship will facilitate his personal growth.

Several other studies carried out by members of the Psychotherapy Research Group of the Wisconsin Psychiatric Institute also have bearing upon the relevance of the Relationship Inventory to the process of positive personality change. A second paper by van der Veen (1962) indicated that the client's perception of Conditions offered by the therapist is inversely related to his degree of disturbance. The less psychologically mature person, then, may be further cut off from the benefits of personal interaction by his own inability to perceive the positive aspects of the interaction. Studies by Truax (1962) and Spotts (1962) both show that the clients whose perceptions of Conditions offered are generally high, or which rise over time, tend to do better in therapy than those whose perceptions of Conditions are generally low or decrease over time. As positive change occurs, perception of meaningful relationships tends to become more favorable also; whether or not a causal relationship exists here, the very correlation of the two aspects of growth may facilitate assessment and even short-term prediciton.

It would seem, from the above discussion, that inclusion of the Relationship Inventory and the sociometric test in the test battery for the present study adds the dimension of _interpersonal_ factors to the more conventional list of individual indicators of positive change. The measurement of simple preferences in liking and esteem among group members and of the perceived presence of therapeutically facilitative factors in two-person interactions, may make possible at least a partial delineation of those specific facets of the group experience which are basic to whatever growth may occur among the individual group members.

Sociometric Test

In any study which addresses itself to social and/or interpersonal variables, it is desirable to have some measure of the manner in which the subjects perceive each other. Any such measure, however, seems to be vulnerable to criticism on one score or another. If the subjects are asked, more or less directly, about their perceptions of each other, there may be conscious or unconscious omissions of quite relevant material. If a less direct method is used, such flaws as experimenter bias and errors of inference are likely to occur.

One of the advantages of a multi-approach study such as the present one is that, in addition to their primary functions, the various instruments may serve as partial validity checks on each other. Thus, if a particular subject has some gross personality problem which might interfere with a valid response to direct questioning, one of the less direct instruments might well pick this up. With this aspect of the research in mind, it was decided to use the direct question methods of sociometry to attack some aspects of interpersonal perception.

Traditional sociometric techniques do not lend themselves easily to statistical analysis. The present authors, in a previous study, attempted to quantify sociometric test results and submit them to statistical test; this proved quite difficult and, in some cases, important relationships could not be tested. Experimental groups cannot be compared as to the relative positiveness or negativeness of their perceptions, inasmuch as the means and variances are necessarily equal between groups. This is also true when one attempt to compare the overall perceptions of one subject with those of another.

Gardner and Thompson (1956) have developed a technique which allows the sociometric subject to use an external reference point, or series of points, in rating his peers. This technique has been published as a sociometric test for elementary school children (1959), but this published test was not suitable either in content or in level of sophistication for the present population. We therefore developed an instrument of our own which utilized the Gardner and Thompson principle but applied it within a slightly different frame of reference.

Garner and Thompson's principle, stated briefly, is this: require the subject to set up an external frame of reference bounded by the people who, in his life experience, represent the extreme examples of the variable in which you are interested. Then ask him to rate his peers according to how they compare with these extremes. In the present study, we were concerned with three major interpersonal-perceptual variables: liking for the other members of the group, respect for the other members of the group, and the degree of comfort which the subject felt in the group. Accordingly, we constructed an instrument which had three parts, one dealing with each variable.

In each part, the subject was asked to write down the initials of the person in his life who most represented and who least represented the quality at issue. Thus, he was first requested to choose the person whom he liked most, of all the people he knew, and the person he liked least. Then he was to choose someone about half-way in between these two. The last step of the reference-establishing phase was to select someone about mid-way between the top and the middle, in terms of his own personal hierarchy of liked and not-liked people, and someone about mid-way between middle and bottom. These selections gave him five specific anchor points. The initials were written on a specially marked sheet, and between each pair of adjacent names was another nonspecified anchor point. In the rating phase of the test, the subject assigned a number between 1 and 9 inclusive (corresponding to the 9 anchor points) to each of the members of his group.

Such a procedure has several advantages over the more traditional sociometric test. First, it allows standard statistical analyses to be performed. Since the subjects are not constrained as to the numbers of the ratings which they may assign the other group members (as is the case when the subject must rank-order them according to a given criteria), means and standard deviations among groups and among individuals are not fixed. The experimental variables built into the design may be examined by means of relatively simple analysis of variance techniques.

With this type of sociometric test the subjects seem to find the task more well-defined. It is as if the reference-establishing phase allows them the opportunity to settle to their own satisfaction just what is the nature of the variable upon which they are rating. Of course, this variable may differ somewhat from subject to subject, since each works independently; but this is also true of more traditional methods. Also, the frame of reference cannot shift during the rating period; it is fixed and specifically anchored at five points along the scale. We can be reasonably assured that each group member rated by a given subject is being rated on the same characteristic or characteristic cluster.

It should perhaps be pointed out that six basic scores per subject--not three--are obtained from the sociometric test. We can measure the way each subject perceived the other members of his group, in terms of liking, respect and comfort; we can also measure the way he is perceived _by_ the other members of his group. In addition, of course, these individual perceptions can be used in various combinations and contrasts, both among themselves and in connection with the other instruments used in the study.

Teacher Ratings

Often in so-called "scientific" studies of human behavior, such an attempt is made to objectify the various measures and criteria that some of the more natural (and meaningful) sources of information are quite overlooked. In the present study, it was felt that the high school teacher, as a person who spends much time with the students and knows them relatively well, might provide such a natural information source. In order to tap this source, a student rating scale was developed and distributed to teachers in each of the three schools involved.

Students were to be rated in four categories: social maturity, self-knowledge, academic maturity, and socioeconomic level. The first three of these seemed to the experiemnters to be of a qualitatively different nature than the last, but it was assumed that the difference in quality of category would not preclude rating the students on the same kind of scale as the others. Ratings in all categories were therefore on a 7-point scale, from "extremely low" to "extremely high." Definitions or descriptions were given to the teachers for neither the rating categories nor the seven anchor points, under the assumption that the intuitive definitions used by the teacher might make for more valid ratings.

In two of the three schools (the parochial and the university high schools) it was possible for a single teacher to rate all of the students from that school. In the public high school, no single teacher could be found who knew all of the students in the sample group, and so several teachers had to be used. These teachers' contacts with the subject were either as homeroom teachers or as American history teachers.

The rating task itself was rather mechanical, as can be seen from the sample rating sheet in Appendix B. The teachers were provided with a rating sheet for each student, the student's name typed at the top, and an addressed envelope in which the finished ratings were to be sealed and left for the experimenter.

Whatever validity an instrument of this sort may have must generally come under the heading of "content validity." To the extent that the definitions of the teachers of "social maturity," "self-knowledge," "academic maturity," and "socioeconomic level" agreed, both among themselves and with the experimenter, the rating instrument did have a high degree of content validity. However, certain indications of lack of agreement, as well as insufficient information on which to base a judgment, have caused us to question both the reliability and the validity of this instrument.

References

Adorno, T. W., et.al. *The Authoritarian Personality*. New York: Harper & Brothers, 1950

Barrett-Lennard, G. T. Dimensions of therapist response as causal factors in therapeutic change. *Psychol. Monogr.*, 1962

Berlin, J. I., and Gendlin, E. T. Autonomic correlations of interaction processes. Mimeographed paper, Wisconsin Psychiatric Institute, 1960

Block, J. A comparison of the forced and unforced Q-Sorting procedures. *Educ. psychol. Measmt.*, 1956, *16*, 481-493

Butler, J. M. and Haigh, G. V. Changes in the relationship between self-concepts and social concepts consequent upon client-centered counseling. In C.R.Rogers & R.F.Dymond (Eds.) *Psychotherapy & Personality Change*. Chicago: Univ. of Chicago Press, 1954, pp. 76-98

Cohn, T. The relation of the F-Scale to intelligence. *J. Soc. Psychol.*, 1957, *46*, 207-217

Comrey, A. A factor analysis of items on the MMPI depression scale. *Educ. psychol. Measmt.*, 1957, 17, 578-585

Dymond, Rosalind F. Adjustment changes over therapy from self sorts. In C.R.Rogers & R.F.Dymond (Eds.), *Psychotherapy and Personality Change*. Chicago: Univ. Chicago Press, 1954, pp. 76-98

Ellis, A. The validity of personality questionnaires. *Psychol. Bull.* 1946, *43*, 385-440

Engel, Mary The stability of the self concept in adolescence. *J. Abnorm. Soc. Psychol.*, 1959, *58*, 211-215

Frank, G.H. Note on the reliability of Q sort data. *Psychol. Rep.*, 1956, *2*, 182

Hart, I. Maternal child-rearing practices and authoritarian ideology. *J. Abnorm. soc. Psychol.*, 1957, *55*, 232-237

Jackson, D. Cognitive energy level, acquiescence, and authoritariansim. *J. soc. Psychol.*, 1959, *49*, 65-69

Lambert, P. Interaction between authoritarian and nonauthoritarian principals and teachers. *Genet. psychol. Monogr.*, 1958, *58*, 163-205

Lewis, Madge K. Counselor prediction and projection in client-centered psychotherapy. Unpublished doctoral dissertation, Univ. of Chicago, 1959.

Lipitz, M. The effects of information on the assessment of attitudes by authoritarians and nonauthoritarians. *J. abnorm. soc. Psychol.*, 1960, *60*, 95-99

McCall. R. J. Face validity in the D scale of the MMPI. *J. clin. Psychol.*, 1958, *14*, 77-80

McClintock, C. G. Personality syndromes and attitude change. *J. Pers.*, 1958, *26*, 479-493

Mozar, R. E. Three versions of the F Scale and performance on the semantic differential. *J. abnorm. soc. Psychol.*, 1960, *60*, 262-265

Rogers, C.R. The necessary and sufficient conditions of therapeutic personality change. *J. consult. Psychol.*, 1957, *21*, 95-103

Rogers, C.R. *On Becoming a Person*. Boston: Houghton Mifflin, 1961.

Rogers, C.R., and Dymond, Rosalind F. (Eds.) *Psychotherapy and Personality Change*. Chicago: University of Chicago Press, 1954

Spotts, J.E. The perception of positive regard by relatively successful and relatively unsuccessful clients. Unpublished research report, Wisconsin Psychiatric Institute, 1962

Stephenson, W. *The study of behavior*. Chicago: University of Chicago Press, 1953

Truax, C.B. The relationship between the patient's perception of the level of therapeutic conditions offered in psychotherapy and constructive personality change. Unpublished research report, Wisconsin Psychiatric Institute, 1962

van der Veen, F. Client process movement and the perception of therapist conditions by clients, therapists and judges. Unpublished research report, Wisconsin Psychiatric Institute, 1962

van der Veen, F. Perceived therapist conditions and degree of disturbance: a comparison of conditions perceived by hospitalized schizophrenic patients and counseling center clients. Unpublished research report, Wisconsin Psychiatric Institute, 1962

Welsh, G.S. and Dahlstrom, W.G. *Basic readings on the MMPI in Psychology and Medicine*. Minneapolis: University of Minnesota Press, 1956

Zuckerman, M. and Oltean, Mary Some relationships between maternal attitude factors and authoritarianism, personality needs, psychopathology, and self acceptance. *Child Development*, 1959, *30*, 27-36

Chapter III

Plan of the Research

Design.--Figure (III-1) presents the design of the study. As can be seen, twelve cells, containing a total of 72 subjects, constitute the experimental sample. In addition, a control group has been observed over a comparable period of time.

The two principal independent variables are congeniality and structure level. It is also possible to isolate effects due to sex, to different counselors, and to different school settings. The latter variable, however, is somewhat confounded with structure level, inasmuch as the sample size did not allow complete nesting within schools. The original design also called for a comparison of groups composed of all boys, all girls, and mixed boys and girls. Again, the number of available subjects did not allow this variable to be studied at the same time as the congeniality variable; instead all groups were composed of three boys and three girls.

	Counselor		Boys	Girls
Public High School	A	C_j	3	3
		U_j	3	3
	B	C_j	3	3
		U_j	3	3
University High School	C	C_j	3	3
		U_j	3	3
	D	C_j	3	3
		U_j	3	3
Catholic High School	E	C_j	3	3
		U_j	3	3
	F	C_j	3	3
		U_j	3	3

CONTROL	C_j	Boys Girls
	U_j	Boys Girls

Legend:

C_j -- congenial groups

U_j -- non-congenial groups

High structure level:	counselors			D and F
Middle	"	"	"	B and C
Low	"	"	"	A and E

Figure III-1. Design of the study, showing schools, counselors, congeniality, and structure level variables.

Subjects.--Subjects for the study were high school juniors drawn from three school settings: a public high school, a university-affiliated high school, and a Catholic high school. The choice of 11th-graders was made in order to utilize subjects who had been in the high school milieu long enough to be thoroughly familiar with and assimilated into it; while at the same time available for follow-up investigation during the following school year. Three schools were used both for practical reasons (it would have been extremely difficult to obtain a sample of the size and characteristics necessary for the study if only one school population were available) and also in an attempt to make more generalizable whatever conclusions might be drawn from the data.

Selection of subjects was on a volunteer basis; within the volunteering population, random assignment was used wherever the dictates of the design allowed. One of the principal investigators of the study visited each of the schools and spoke to all of the 11th-grade students (in two of the three schools, this was done in several smaller groups; in the university high school it was possible to address all the students at once). The project was described as an attempt to learn more about the way in which groups of teenagers discuss things in small groups. Students were told that if they volunteered and were selected, they would meet in a small group setting twice weekly for a total of fourteen meetings. The meetings were to center around the general topic, "How do we see people?" It was emphasized that volunteering carried with it a commitment to continue through the whole series of fourteen meetings. At no time was the idea of varying either congeniality or structure level mentioned, nor was there any more specific description of the aims of the study than the rather ambiguous statement given above.

After this general introduction, sign-up sheets were passed around. Students who volunteered were asked to indicate other 11th-graders in their school with whom they would like to meet in the small groups; no restriction was made on the number of selections a student was allowed to make.

In the public and parochial high schools about one-fourth to one-third of the junior class students indicated a willingness to participate in the discussions. In the university high school (a much smaller school) about 90 percent of the total volunteered to take part.

The first step after the sign-up sheets had been collected was to arrange them into congenial and non-congenial groups. This proved to be a much more difficult task than was anticipated. The minimum criterion for a congenial group was that each person be chosen by at least one other in the group, and that the choices not be so arranged as to split the group into two or three smaller units between which existed no choice bonds. Beyond this minimum requirement, congenial groups were so chosen as to contain as many mutual choices as possible, and as many total choice bonds as possible. Non-congenial groups were allowed neither mutual nor single-directional choices. In other words, no member of a non-congenial group was allowed to have indicated any other member of that group as a preferred co-member.

It soon became apparent, in working with the various lists and tentative groupings, that the great majority of mutual choices (and thus of congenial-group members) were among people who were also chosen by classmates who were themselves more or less social isolates. Therefore, in order to meet the criteria for a congenial group, we were forced to fill it with students who were chosen by many others. The reverse was true for the non-congenial groups: these tended to be filled by students whom no one had chosen as a preferred co-member of a group.

The stringency of the criteria for group membership also made it quickly apparent that virtually all of the original sample would be needed to fill the experimental cells of the design. It was therefore decided that rather than cut down on the number of variables to be examined, the control group would be set aside for that school year; a group as nearly as possible isomorphic to the orignial would be selected the following year and used as a control.

Procedure.--Once the two congenial and two non-congenial groups had been selected from each school, they were randomly assigned to one of the counselors who would work in that school. Assignment of counselors to schools had previously been made in such a way as to balance the structure level as nearly as possible among schools.

Structure levels have been defined in a previous section (see Chapter II).

Counselors were assigned to one of the three structure level conditions on the basis of their past experience and of their present preferences. Every attempt was made to place each counselor in that setting in which he would feel most comfortable and work most effectively. The two counselors working under the low structure condition were both doctoral students whose general counseling orientation was in a nondirective setting. The two high-structure level counselors had both been classroom teachers at one time, and felt most comfortable when a certain degree of structure was available. In the middle structure class were a youth worker from one of the campus churches, and a graduate student with considerable training in psychology and counseling.

After all group assignments were completed, students were informed that they had been selected and asked to meet with their group leader the following Saturday in order to select a regular meeting time and also to fill out some questionnaires. The "questionnaires" which they filled out were the Q sort, the California F scale, and the WLP. Some also filled out the MMPI.

Subsequent meetings were held with the group leader, under the conditions stipulated by the structure level within which each group operated, for a total of 14 bi-weekly meetings. Tape recordings were made of each group session. Following the first of these regular meetings, students filled out the Relationship Inventory and the sociometric test. Following meetings 13 and 14 of the series, the Q sort, F scale, WLP, Relationship Inventory, and sociometric test were re-taken. Some time after the study had been completed, teacher ratings were obtained on the students involved (see Chapter IV).

Follow-up data, utilizing all psychometric instruments with the exception of the MMPI, were obtained approximately __ months after the completion of the study and again approximately one year later.

Chapter IV

Tape Analysis Procedures

(to be completed)

Chapter V

Results and Discussion

While the attrition rate in most of the cells was quite small (in nine of the 12 cells no subjects were lost) there was almost total attrition in one cell. The nearly inevitable problem of missing bits of data, due to absences at time of testing or invalidation of protocols for one reason or another, also had to be dealt with.

Statistical techniques do exist for handling analysis of variance designs with unequal cell frequencies. We had written for the study a computer program for the case of an n x m unequal cells design--to the best of our knowledge, the most general unequal cell program available at present--and are in the process of extending this program to handle an n x m x r design where r represents a repeated or correlated variable. However, preliminary tests with this unequal cell technique suggested some fundamental weaknesses, primarily in the power of the tests for interaction effects.

For this reason, all of the analyses to be reported on the following pages have involved standard, equal-cell frequency designs. Cell n's were dropped to 5 for all analyses. Wherever necessary, absence or nonvalidity of a particular piece of data dictated what should be dropped; elsewhere the dropping was by random selection. In those rare instances in which more than one piece of data was absent from a given cell, the necessary fifth element was extrapolated from the other cell entries and from that subject's behavior on related instruments.

Due to the virtual loss of one cell for post-treatment testing, certain of the variables had to be collapsed over many of the analyses. These collapsings will be apparent in the discussion of each of the analyses.

In view of the literature about the California F scale (see Chapter II) and of the F Scale results in the authors' previous study of small group interaction, it was expected that F scale differences would occur among the subjects of the present study. These differences do show up, both initially among the different groups, and over time. Figures V.1 and V.2 show the analysis of variance results when the F scale scores are compared among levels of congeniality and structure, over the period of the group interactions. As can be seen, the difference between the congenial and non-congenial groups is highly significant, with the congenial groups scoring lower on the F scale. The pre-post variable approaches significance, with the post-treatment scores tending to be lower ($p < .10$). The interactions between trials and structure, though not significant at even the .10 level, is nevertheless interesting as a trend, and the difference among the groups does at least raise the question of "real" as opposed to chance variation ($F = 2.70$; 2.85 needed for significance at $p < .10$).

The congenial/non-congenial overall difference is one of the first indications of a confounding in the congeniality variable. Due to the method of selection (see Chapter III), subjects in the "congenial" groups tended to be the more popular, well-liked students in the school; while subjects in the "non-congenial" groups tended more toward the social isolate end of the scale. It is difficult to imagine how such a variable could have been avoided in a workable design; be that as it may, it is necesaary to consider the congeniality factor to be strongly weighted--if not overshadowed--by an element which might be variously described as popularity, leadership, socialibility, or the like.

The student who is well-liked by his peers tends, in general, to be at least as bright as average, often outstanding in one or more areas, and to be outgoing and psychologically well-balanced and mature. The F scale has been shown to correlate with both intelligence and academic achievement; e.g., bright subjects tend to score lower on a measure of authoritarianism. There also seems to be a relationship between psychological health and authoritarianism, as measured by the F scale, with more healthy subjects appearing less rigid on the test. With these relationships in mind, it is not surprising that the "congenial" subjects in the present study should score lower on the F scale than the "non-congenial" subjects.

The pre-post difference is encouraging, though not sufficient for claims of psychotherapeutic success in and of itself. If the trend does in fact represent a change as a result of the experimental procedures, it would seem that group counseling did reduce the subjects' level of authoritarian rigidity.

The trend of the trials-by-structure interaction, though not statistically significant, is worthy of comment inasmuch as it supports, in an almost classic fashion, some of the Rogerian suppositions about group functioning. Both of the structure level groups tend to decrease their authoritarian rigidity scores; but the low structure group shows much more change. This tendency allows us to hypothesize that the nondirective group setting, in which group members are encouraged to proceed in whatever direction and at whatever speed is confortable, may be more conducive to positive change in this area.

R.I. Analysis

As was described in the plan of the research (Chapter III), each subject in the study filled out three Relationship Inventories at each testing period. The first R.I. was answered with regard to the person in the subject's group whom he liked most; the second with regard to the person whom he liked least, and the third with regard to the group as a whole--the group cum group, as it were. These Most, Least, and Group administrations form effective subtests of the R.I. as a whole, with the four subscales "nested" within each subtest. Treating the instrument in this fashion yields a 5-way analysis of variance model (see Figure V.3). The analysis of variance sheds light on some of the general factors involved in the overall functioning and internal structures of the groups, as well as the dynamic effects of structure and treatment.

Figure V.4 contains the source table for the overall analysis. One of the most surprising results is the lack of significant differences between the Congenial and the Non-congenial groups. In view of the supposition that the R.I. is at least partly dependent upon the degree of positive feeling on the part of the respondee toward the person about whom he fills out the test (born out by the contrast between scores on Most-liked person and on Least-liked person), this lack of significance supports our suspicions that the Congenial/Non-congenial variable might be more accurately conceptualized as a "popular student/less-popular student" variable.

The significant M-L-G differences ($F = 152.8$; $p < .005$) follow the expected pattern: Most-liked person is rated highest, Least-liked lowest, with the Group in the middle. The main effect for scales is also highly significant ($F = 170.9$; $p < .005$).with Positive Regard (R) and Congruence (C) scoring higher than do Empathy (E) and Unconditionality of Regard (U). Evidently among students of this age it is easier to see (or to project from one's own feelings) a general liking for and honesty with oneself, on the part of a peer, than to attribute to him a more altruistic unconditional regard for and empathy with oneself. An examination of the individual items on the test also suggests that the R and C subscale item might appear somewhat less "goody goody" or "queer" to the average adolescent than those of the E and U subscales.

The interaction between M-L-G and scales is significant ($F= 14.7$; $p < .005$) with the relationship among the four scales differing between Most and Least-liked persons (Figure V.5). For the Least-liked person, Congruence is highest, with Positive Regard second; this is reversed in the Most category. Evidently the person liked most in the group is seen with liking the respondee as a most outstanding trait. The person liked least, however, is seen to like the respondee less but to be more honest with him (in comparison to the "liking" or Positive Regard). Another aspect of these differences shows itself in the wider spread among the R scores of the three groups, compared to the relative bunching of the C and U scores. It is apparently the overall perceived regard of the Other for the respondee which most differentiates the Most from the Least-liked person within the respondee's perceptual field.

The Congeniality variable becomes apparent as a discriminating factor in its interaction with scales ($F = 4.13$; $p < .01$). The non-congenial groups tend to vary less in their scores from scale to scale, among the R, E, and C scales. It might be hypothesized that these students are less able to differentiate the several qualities which they are asked to evaluate in their peers, seeing instead a more unitary or negative relationship.

Considering overall perceptions once more (i.e., collapsing across Scales) we find a significant interaction among Congeniality and the M-L-G division ($F = 7.29$; $p < .005$). Examination of the relationships among group means indicates that the interaction effect occurs in the ratings of the Group as a whole. The Most and Least scores fall about as one would expect, with Most scoring higher than Least and Congenial higher than Non-congenial. In the Group scores, however, the scores for Congenial and Non-congenial almost coincide, with the Non-congenial slightly higher. Evidently the group as a whole can be seen in a relatively positive light, even when rankings of individuals in the group are less favorable than those in other groups (see Figure V.6).

Turning now to changes over time, we see a trend toward overall improvement in the main effect for Trials ($F = 3.1$; $p < .10$). Looking at the interaction between Trials and Structure, also bordering on significant ($F = 2.95$; $p < .10$), it is apparent that the positive perception changes occurred in the high structure groups; the means of the low structure groups are identical in the pre- and post-treatment trials.

An examination of the Structure x Congeniality x Trials interactions ($F = 4.92$; $p < .05$) yields an even more interesting set of relationships, as shown in Figure V.7. Here we see the Congeniality variable acting in opposite directions in the high-structure and the low-structure groups. Under high-structure conditions, the congenial groups increase their scores much more than do the non-congenial groups. Under low structure conditions, the congenial groups' scores actually decrease slightly, while the non-congenial groups increase more than they did under high structure. We can only speculate, of course, as to the dynamics of these contrasts: do the congenial students actually become less fond of each other under more permissive conditions? Is the experience of free, non-threatening group interaction so novel to the non-congenial groups that they thrive on lack of structure? Are the bright, popular students assigned to the congenial groups so used to being leaders and innovators that they perhaps relish the novelty of "following" in the high-structure situation and respond particularly well to it?

While we cannot answer the "why" questions on the basis of the Relationship Inventory results, we can make some generalizations about the overall situation. It seems rather clear that, over the period of the study, students' perceptions of each other did tend to become more positive. This seems to be as true for the Most-liked person as for the Least-liked person--in sharp contrast to the authors' findings with college students in similar groups, where Least-liked person was perceived even more negatively at the end of the study. Also, the structure variable seems to have had quite different effects upon the Congenial and the Non-congenial groups, with high structure facilitating the Congenial but not the Non-congenial. Finally, the student seems to rate the other's perception of himself highest in a sort of general category of "he likes me," and lowest in a more altruistic sense of "he likes me no matter what I do."

Q Sort

One of the pecularities of the Q sort as a psychometric tool is that one administration cannot, by itself, yield a quantifiable score. The Q sort is essentially a normalized ranking technique: as such, sums, means, and variances of all possible distributions are the same. In order to obtain quantitative score one sort must be compared (most usually by means of correlation) with another.

As was described in Chapter III, each student in the present design was asked to perform two kinds of Q sorts. First he sorted according to the kind of person he thought himself to be (self sort) and then as he would like or wish to be (ideal sort). These two sorts, administered pre- and post-treatment, yielded, for each subject, six comparisons: self-ideal pre, self-ideal post, self pre-self post, ideal pre-ideal post, self pre-ideal post, and ideal pre-self post. Only the first four of these comparisons have been treated in the present analysis.

In addition, comparisons were made of all four subject sorts with the Expert sort (see Chapter II).

The distribution of the statistic r. (e.g., of possible values of the product-moment correlation coefficient) is not normal. Thus parametric statistical analyses, such as the t test and the analysis of variance, which assume normality of distribution, cannot legitimately be used to evaluate differences among groups of r.'s. However, Fisher* has shown that the z' transformation can be used to normalize the distribution of r. For this reason it was necessary, as a preliminary to statistical analysis, to convert all of the correlation coefficients to z' scores. The z' scores were then transformed to integers correct to four significant digits.

The two self-ideal comparisons, pre- and post-treatment, were subjected to a 2 x 2 x 2 analysis of variance, using structure level, congeniality, and trials as the independent variables. Figures V.8 and V.9 show the source and mean tables for this analysis.

There is, in this analysis, a significant difference overall between the high-structure and the low-structure group. This is difficult to understand, inasmuch as "structure," a treatment variable, would be expected to take effect over time, rather than as a main effect. However, if one is allowed to speculate, it could be hypothesized that the orientation of the groups, obvious after even the first session, may have influenced test-taking attitudes on the part of the subjects. The high-structure groups show more correspondence or agreement between self and ideal sorts. The structure of the group may have focussed attention on psychological and personal attributes and created attitudes of self-awareness and/or defensiveness with regard to the kind of areas sampled by the Q sort; this could conceivably raise the self-ideal correlation.

---

*Fisher, R. A. On the "probable error" of a coefficient of correlation. Metron, 1921, 1, Part 4, 1-32.

The overall difference between trials is also significant, with the post-treatment correlations higher. The general affect of the group interactions seems to have been to raise the correspondence between perception of self and of ideal self.

Figures V.10, V.11, and V.12 show the analysis between the self-self and the ideal-ideal correlations. Again we see the overall higher correlations obtained by the high structure group, although in this analysis the F ratio misses significance at the .05 level. There is a nonsignificant trend for the overall correlations among the congenial groups to be higher than those among the non-congenial, suggesting that there may have been less change over time in the former (higher correlations indicate greater similarity between sorts). As would be expected, the ideal-ideal correlations were significantly higher than the self-self.

The most interesting aspect of this analysis is the highly significant interaction between structure and the self/ideal variable. As can be seen in Figure V.12, the self-self correlations in the low structure groups are much lower than any others. Indeed, it is safe to assume that it is this "straggler" group which creates the significant F ratio for the interaction. Since the low correlation indicates greater differences between the two sorts, we can conclude that the majority of the changes over time occurred in the self sorts of the low structure groups; the ideal sorts of the low structure groups, and both self and ideal sorts of the high structure groups, seem to have changed only slightly.

These differences are reflected in a slightly different way in the analyses of the self-Expert and ideal-Expert correlations (Figures V.13, V.14, V.15, and V.16). For the ideal-Expert analysis, there is no significant overall difference between the high- and low-structure groups, and no pre-post differences. The self-Expert analysis tells another story, with the highly significant between-structure F ratio and a significantly higher post-treatment group mean.

The significant structure x trials interaction in the ideal-Expert analysis sheds still more light on the relationships involved. Here we see that while the similarity of the low structure groups' ideal sorts to that of the Expert increased over time, that of the high structure groups tended to decrease.

It is also interesting to note that, for the ideal-Expert comparisons, the congenial groups tend to score higher--that is, their perceptions of their ideal self tend to be more like the Expert sort than do those of the non-congenial groups. For the self-Expert correlations, however, the relationship is reversed; here the non-congenial groups perceived themselves in a more similar fashion to the Expert than did the congenial groups.

Finally, while not quite reaching significance, the second-order interaction between structure, congeniality, and trials in the ideal-Expert analysis is somewhat enlightening. As Figure V.16 shows, each set of subgroups tended to increase its similarity to the Expert sort over time, except for the high structure non-congenial groups. Here the tendency is sharply in the opposite direction.

In summary, one may view the Q sort as yielding rather strong evidence that positive changes did occur among some of our subjects, but that negative ones may have occurred elsewhere. The perception of self tended, overall, to improve, as did similarity between perceived self and ideal self. The low structure groups show more consistency, with the general trend always in a positive direction. The high-structure groups, while reflecting positive changes in perception of self, show congeniality level differences in changes in perception of ideal self. There is no real evidence that either the congenial or the non-congenial groups were in general more facilitative.

It was originally hypothesized that students who liked each other would show more similarity in their self-sorts than would those who did not express a liking for each other. To test this hypothesis, pairs of students were chosen within groups who rated each other high on the sociometric instrument (indicating high liking or mutual regard) and who rated each other low on the instrument (indicating low liking or mutual regard). Only pairs in which the ratings were mutually high or low were chosen.

Correlations were then run between the self-sorts of each of these pairs, and the z' conversions of the correlations were compared between the high regard and the low regard pairs on both pre-treatment and post-treatment data. Over the entire sample, no values of t significant at or below the .05 level were observed, and the hypothesis was not upheld. However, the differences between high and low mutual regard pairs in the post-treatment comparison did approach significance with $p < .10$ (see Table V.17).

Before discarding this line of inquiry entirely, a number of factors should be pointed out. First, the rather stringent criteria for inclusion in a high- or low-regard group meant that the n's for the comparisons were quite small, and that the majority of the subjects in the sample did not appear. Inasmuch as the qualification of one subject pair for inclusion did affect the chances that any other subject pair would qualify (the criterion for inclusion being relative, within the sociometric instrument data, rather than absolute), the assumptions of independence underlying the t test may have been seriously endangered. Secondly, the pooling of subject pairs from all groups greatly increased the error variance and thus reduced the power of the t test considerably. Within group comparisons would have eliminated this source of type II error, but the n was too small to carry out such an analysis.

Finally, the existence of invalid sorts and the problem of missing data were quite crucial in this analysis. In the comparison of pre-treatment pairs, one-half of the possible low mutual regard pairs had to be discarded because of missing or invalid data. More of the possible high mutual regard pairs had to be discarded for this reason. It seems probable that careless sorting--leading to invalid data--or not showing up for the testing period at all was a systematic factor which mitigated against accurate testing of the hypothesis.

In conclusion, then, while the present data clearly failed to support the hypothesis of difference between the two groups, it is suggested that the available data did not provide an adequate test, and that replication is needed before any conclusions can be drawn.

Sociogram Analysis

Due to the nonranked nature of the sociometric instrument used in the study, it was possible to treat sociometric test scores in the same fashion as any other set of parametric data. These scores were organized in two ways: one, the average scores given by each individual to the other members of his group; and two, the average scores received by each individual from the members of the group. It was necessary to use averages rather than sums because the attrition factor resulted in unequal n's among groups, and also some of the subjects joined their respective groups late and thus were unavailable for rating on the pre-treatment testing.

Considering the scores Given and the scores Received as subtests of the sociometric test as a whole, and each as containing the three subscales of Liking, Respect, and Comfort, yields a 2 x 2 x 3 x 2 x 2 repeated measures design (see Figure V.18). The source table for the analysis is presented in Figure V.19.

In this analysis, as in the analysis of Q sort results, we find a significant difference between overall high-and low-structure groups, with the low structure groups rating and being rated by their peers more positively. Again, this is difficult to understand inasmuch as structure is a treatment rather than an assigned variable. It is possible that here too the atmosphere of the group was apparent from its inception, and this created different test-taking sets. Another possibility is that the changes over time, occuring as they did in only one group (see below), were sufficiently large to create an overall difference as well as a significant interaction.

The significant difference between Congenial and Non-congenial groups ($F = 27.5$; $p < .01$) was in the expected direction, with congenial groups scoring higher. As before, of course, interpretation of this effect is confused because of the nature of the group: we would expect this sort of result both in comparisons of groups of people who liked each other and groups who didn't, and in groups of popular, likeable people vs. groups of retiring, not-so-popular people (see Figure V.20).

The hypotheses regarding the repeated measures of this (and other) analysis having been nondirectional, it has been necessary to use a two-tailed test of significance on the data. For the sociometric data, this has often had the effect of reducing the level of significance to <.10, not generally considered within the acceptable range. Here, as elsewhere, however, these interactions will be reported; it is felt that their value as hypothesis-generating trends justifies their inclusion.

There is a difference among the Scales, with Respect lowest and Liking slightly higher than Comfort. This is in agreement with our subjective appraisal of the reactions of participating students: it seemed to be difficult for them to think of each other in terms of "respect," whereas "liking" was an easy and familiar concept and "comfortable to be with" only slightly less so.

Trials show significant differences but in an unexpected direction: the post-treatment ratings are lower than the rankings when the subjects knew each other less well. Apparently some aspect of the group interaction had a negative effect upon the students' perceptions of each other. Looking at the Structure by Trials interaction, also significant at the .10 level, we gain more insight into the nature of these changes: the low structure groups changed only negligibly, while the real decrease came among the high structure subjects (see Figure V.21). We can trace the pre-post relationship a bit further by examing the Structure x Scales x Trials interaction, where we find significance at the .05 level. Here we see no real difference between high structure and low structure groups on either the Liking or the Comfort scale; but where the low structure groups increase their scores on Respect over time, Respect scores in the high structure groups go down strikingly (Figure V.22).

These results are curious both in and of themselves, since on the face of it there would appear to be nothing in the high-structure procedure which would tend to decrease the mutual esteem of the members; and also in relation to the Q sort and Relationship Inventory, both of which indicated that the high structure group is facilitative of positive perceptual changes. In the R.I., of course, only the two extremes of the group are being considered; moreover, any comparison must be made within the group. The sociometric test considers the whole group on the basis of external criteria. It is possible that both the R.I. and the sociometric results can be attributed to a heightened awareness and accuracy of interpersonal perceptions, both within and without the treatment group. Clearly, however, this is a highly ambiguous area and one which must be explored more fully before even tentative conclusions can be drawn.

Relationships between the Congenial and the Non-congenial groups can also be seen more clearly in terms of their interactions with the repeated measures. Surprisingly, changes over time in the two types of groups are almost identical. There is a trend, significant at the .10 level, in interaction of Congeniality with Scales; and the interaction with the Given/Received variable is significant ($p < .05$). With regard to Scales, while Liking is high in both groups, Respect is low only in the Congenial groups. The lowest mean score of all is the mean for Comfort in Non-congenial groups. Whether this is indeed due to the subjects' perception of each other, or to some inability to be really comfortable in any social situation, is not clear (Figure V.23).

Non-congenial groups scored lower than Congenial groups in both scores given to others and scores received by others. However, considering the Congenial and the Non-congenial groups separately, we see that among the Congenials the scores Given were higher than the scores Received, while among the Non-congenials the reverse was true. In other words, the Non-congenial subjects tended to see their peers less positively than they themselves were seen, while the Congenial subjects gave higher ratings to their fellow group members than they themselves received.

It should be emphasized again that none of these repeated measure main effects and interactions reaches an acceptable level of significance. Any speculation about their meaning must be considered just that: speculation. Nevertheless, the patterns are provocative. Whether artifact of design or of instrument, or reflective of real factors operating in the study, further investigation of these interpersonal variables could hardly fail to be productive.

WLP

Prior to examining the WLP data in the context of the experimental design proper, it was decided to run inter-item correlations within each subscale of the instrument. Inasmuch as this study provided the first real test of the WLP as a research tool, such a check would give further evidence of the cohesiveness of the subscales and, indirectly, of the validity of the instrument itself. Figure V.24 shows the inter-item correlation matrices for each of the eight subscales. As can be seen, items within each subscale are for the most part quite strongly related. The single exception is Scale S-A, Similarity to Adults. Examination of the 5 items in this subscale suggests that two factors are operating here: one, the self-perceived similarity of the subject's behavior to a generalized "adult" behavior, and two, the subject's approval of the actual adult behaviors which he sees.

Figure V.25 shows the analysis of variance for the WLP scores on the eight subscales. It is apparent that the amount of useful information is quite small. None of the main effects with the exception of scales is significant, nor are there any discernable trends. Scales is significant at the .01 level, with Confidence lowest; Basic Values and Trust high.

Bearing in mind the large "regional" differences (where "region" was highly confounded with type of school) in the original normative sample, it seemed possible that the differences in type of schools sampled in the research population might have contributed excessively to within-cell variance in the general analysis described above. For this reason, another analysis was carried out, this time using type of school as a main effect, and collapsing across treatment groups. It would have been more appropriate, of course, to use both treatments and schools; attrition factors made such an analysis impossible.

The source table for this second analysis is presented in Table V.27. It is obvious that the "hunch" about the effect of the different kinds of schools did not bear fruit. Not only did the main effect for schools fail to reveal significant differences, but almost none of the interactions were significant. Only the Scales main effect and the Schools x Scales interaction reached acceptable levels of significance. The significance of Scales could be predicted on the basis of the previous analysis. Schools x Scales reveals a rather interesting pattern (see Figure V.28), but the relatively small order of magnitude of the differences in the group means suggests that the significance of the interaction probably is of greater relevance to an understanding of the instrument itself than to support or rejection of our research hypotheses.

On five of the eight scales, the students from the university-associated high school had the lowest scores. On only one scale, Confidence, did these students have the highest scores. It is interesting to note, in this context, that the university high students tended to be, in general, a rather flip, self-assured, cynical group of young people. Differences between the parochial and the public high school students were quite slight on most scales; the exceptions being scale Similarity to Adults and Basic Values, on which the parochial students scored high, and Pity and Blame, on which the public high school students scored high.

It is difficult to explain the failure of the WLP to reflect differences in and/or changes among the various groups of this study. If one accepts the findings of the normative study, the instrument itself would seem sufficiently sensitive to attitudes and values in this area--and this is the specific area with which two of the three treatment groups were to concern themselves. It is true that the normative study did not take into consideration test-retest effects; perhaps the instrument is not sensitive to changes, or at least to short-term changes. It is also possible that the smaller and more intimate nature of the test-taking groups in the present study tended to inhibit openness and frankness in this test. In view of the highly significant results obtained using other instruments, one cannot justify a conclusion of no real differences or changes in the groups on the basis of the WLP results. At present, all that can be concluded is that the changes were sufficiently slight and/or subtle that the instrument failed to pick them up.

Teacher Ratings Analysis

Teacher ratings were obtained on all students on four variables: social maturity, self knowledge, academic maturity, and socioeconomic status. In order to avoid creating special treatment and/or interest in the experimental subjects on the part of the teachers, which might have influenced both ratings and possibly behavioral or perceptual changes on the part of the students, teacher ratings were not obtained along with the other pre-treatment data. Rating sheets were distributed to teachers only after all other data had been collected. Teachers were asked to rate a student only if they had had him in a class for at least a year.

Both the small range of the ratings assigned by the teachers, and the strong probability of individual differences among teachers doing the ratings, made questionable any statistical procedures using as raw data the numerical values of the ratings. Instead, students were divided into high and low groups on the basis of each of the separate rating scales, and on the basis of the overall ratings received.

The first three rating categories (Self-Knowledge, Academic Maturity, and Social Maturity) were compared among treatment and congeniality groups. Chi-square analyses reveal no significant differences in teacher ratings among the three treatment groups. Comparing congenial with non-congenial groups, however, we find that two of the three rating categories yield significant differences (Self-Knowledge, $X^2 = 4.42$, $p < .05$; Academic Maturity, $X^2 = 13.39$, $p < .01$), and the overall differences were also significant ($X^2 = 9.22$, $p < .01$). This tends to support our notion that the congeniality variable was confounded by a more fundamental difference in the groups. In each of the three categories examined, more of the low ratings went to students in non-congenial groups, and more of the high to students in congenial groups.

Using the overall teacher ratings to divide the students into a high and a low group, comparisons of high rated and low rated subjects were made using F scale and WLP scores as the dependent variables. There are no differences bewteen the two groups on any of the pre-treatment measures; however, some significant differences do occur on the post-treatment measures.

On the F scale, post-treatment, the low-rated students score significantly higher than the high-rated students ($t = 2.95$; $p < .01$). On the WLP post-test, two of the eight scales reveal significant differences between high-rated and low-rated students, with the high-rated group scoring higher ($t = 3.01$; $p < .01$; $t = 2.48$, $p < .02$). The two significant scales are Pity and Blame, and Giving and Taking. In four of the remaining six scales, the high-rated group's scores are higher, but at a non-significant level. The two exceptions are Basic Values, on which both groups are the same, and Confidence, on which the low-rated group scores slightly higher. Again, scale C seems to be consistent in reversing the trends shown by the rest of the WLP.

It was expected that there would be a significant relationship between socioeconomic level and interpersonal attitudes, and between academic attitudes and interpersonal attitudes. This relationship was not conceptualized as a linear one; rather, we hypothesized that interpersonal attitudes would approach an optimal level at the middle range of socioeconomic status and of academic maturity. For this reason, a product-moment correlation procedure was not applicable, and the correlation ratio became the statistic of choice. However, when correlation ratios were run on the relevant variables (teacher ratings vs. Relationship Inventory scores) there were no significant relationships. Indeed, charting the data revealed no groupings or trends, but rather a wide scattering of the data and very slight differences of scores between the different rating categories.

The absence of relationship between these two sets of variables is so apparent that one is tempted to conclude that in our sample, interpersonal attitudes are not affected by socioeconomic status or academic attitudes. The nature of the rating task, as perceived by the teachers who did the rating, suggests another possible explanation. These teachers reported extreme uncertainty and/or lack of information with regard to some of the students whom they rated. Moreover, discussions with the teachers indicated to the experimenters the very real possibility of differing interpretations of the rating categories among the several teachers involved. In sum, it would seem that the teacher ratings were at best a rather unreliable indicator of the variables in which we are interested.

In the original conceptualization of this study, a series of hypotheses were set forth to be tested. As is often the case in research involving a relatively large number of variables, many new aspects and possibilities have occurred to the authors during the course of the project. In order to facilitate the following through of our original lines of thought, however, this section presents a discussion of the overall study in terms of the original hypotheses.

A. Interpersonal attitudes do vary among students.

  1. High authoritarian-attitude students will tend to score lower on initial interpersonal attitude measures than will low authoritarian-attitude students.
  2. The relationship between interpersonal attitudes and academic/socioeconomic factors will approach a normal curve, with higher-scoring students (on initial interpersonal attitude measures) tending toward the middle of the academic/socioeconomic distribution.
  3. Girls will score higher on initial interpersonal attitude measures than will boys.

That there was wide variation in interpersonal attitudes among the students sampled in the study is beyond question. The somewhat disconcerting initial differences between congenial and non-congenial groups alone indicate this to be the case. Relating the differences to specifically identifiable factors, however, is more difficult.

Our reasoning with regard to the initial congeniality level differences has been presented in a previous section. No tests were included in the study which might shed light on the "popularity" factor or indicate the kinds of personality/ability differences which might exist between popular and nonpopular students. We are thus unable to verify our suppositions with regard to the effects of the selection method used.

Correlations were run between F scale scores and Relationship Inventory, Q sort, and WLP scores. These were run across all subjects for whom complete data were available, and on both pre- and post-treatment data. With regard to hypothesis A, we are primarily interested in the pre-treatment relationships. Without exception, correlations between the F scale and the interpersonal variables were nonsignificant; indeed, the absolute value of r. was less than .100 on 12 of the 23 pre-treatment comparisons and on 15 of the 23 post-treatment comparisons. Clearly, we cannot reject the null hypothesis related to sub-hypotheses 1, and must conclude that we were not able to demonstrate a relationship between authoritarian rigidity and interpersonal attitudes.

As was reported in the teacher ratings sections of the previous section, none of the correlation ratios computed for ratings with the interpersonal measures indicated a significant relationship. Thus the second sub-hypothesis above was not supported by the data: neither academic nor socioeconomic factors, as reported by teachers, could be shown to relate to interpersonal attitudes. However, a significant correlation ratio was found between the academic maturity ratings and the F scale scores ($F = 4.4$; $p < .01$). While this pattern is suggestive of an interesting relationship between the two variables, extreme caution should be exercised in its interpretation. Both the questionable reliability and validity of the ratings, and the fact that, out of a possible 30-some correlations, only one reaches significance, indicate the high probability of statistical or sampling artifact being responsible for the observed relationship.

In order to test the third hypothesis of the set, t tests were run between boys and girls over the pre- and post-treatment scores of the sociometric test, the Relationship Inventory, and the WLP. There were no significant differences between boys and girls on either the sociometric test or the Relationship Inventory. However, significant differences did show up on the eight subscales of the WLP. These differences were in every case greater for the post-treatment measures than for the pre-treatment; and on the post-treatment tests, girls scored higher than boys on all eight subscales (scales A, BV, T/m, and BN significant with $p < .05$. On the pre-treatment measures only scale A was significant; however, girls again scored higher on all subscales except G/T and C. From this we can conclude that there are differences between sexes on those aspects of interpersonal attitudes measured by the WLP, that in general girls respond at a more mature level, and that this tendency increases of time and/or treatment.

B. Positive interpersonal attitude change can be facilitated.

1. Positive changes in interpersonal attitudes will occur to a significantly greater degree in treatment groups than in a control group.

(To be completed.)

Hypothesis set C was concerned with changes in interpersonal attitudes over time as a function of group structure and of group congeniality.

C. The degree of positive interpersonal attitude change will vary with more effective in producing positive change than will a highly structured or a free-discussion presentation mode.

1. A moderately-structured presentation mode of discussion will be more effective in producing positive change than will a highly structured or a free-discussion presentation mode.
2. High-congeniality groups will show greater positive attitude change than will low-congeniality groups.

Due to attrition within several of the groups, hypothesis 1 was impossible to test: attrition in the moddle-structure group was approximately twice that of attrition in the other two structure levels. This fact in itself may lead to some conclusions as to the effectiveness of this level of structuring; it could be argued that such a high dropout rate reflects a less meaningful and/or useful experience for the subjects. However, it is the authors' opinion that such a conclusion would be unwarranted; with the small number of counselors, and the widely differing school populations from which the samples were drawn, it is highly likely that a number of interacting factors contributed to the differing attrition rates. It was possible to drop the middle structure group from most of the analyses and to test differences between high and low structures. It is this comparison which is reported below.

In general, the psychometric data produced disappointingly little with regard to these hypotheses. Congeniality/Non-congeniality change differences reached borderline significance only on one measure, the Q sort. Here the trend was in the opposite direction from that which was hypothesized: non-congenial groups tended to show more change than did congenial ($p < .10$; see Table V.10).

The Q sort scores did differentiate between high and low structure groups; although the directionality of the changes was somewhat obscured, the low structure group seem to exhibit more consistently positive change ($p < .05$; see Table V.14). The Relationship Inventory also showed this trend at a borderline level of significance ($p < .10$; see Table V.4). Neither the WLP nor the F scale upheld the hypothesis. The WLP showed no significant differences at all, while the F scale scores suggested that the low structure groups did slightly better than the high ($F = 2.7$, n.s.). On the sociometric test, the general tendency was for scores to go down over time rather than up; the trials x structure interaction indicated that this change occurred only in the high-structure groups, with the low-structure groups holding at about the same level ($p < .10$; see Figure V.21).

Only one measure, then, the Relationship Inventory, tends to support the hypothesis that the high structure groups were superior in facilitating positive attitude change. The others either fail to support this hypothesis or offer some evidence in the opposite direction. While the several measures do, to be sure, concern themselves with differing aspects of the interpersonal situation, this conflict in results--all at a borderline level of significance--cannot in any sense be interpreted as supporting any consistent hypothesis of the superiority of a high level of structuring.

The second hypothesis was also refuted; the only measure which showed any difference at all between congenial and non-congenial groups in terms of change suggested that it was the non-congenial groups which were more facilitative.

Hypothesis set D is concerned with the relationships between changes in interpersonal attitudes and other attitudinal and behavioral changes.

D. Interpersonal attitude change will be associated with other global changes.
    1. Individuals showing greater positive interpersonal attitude change will also show greater positive personality change.
    2. Individuals whose interpersonal attitudes change positively will tend to become less authoritarian.
    3. As interpersonal attitudes change positively, teacher behavior ratings and academic achievement will also show improvement.

Sub-hypothesis 1 will be tested by the relationship between change on the several interpersonal attitude measures (WLP, RI, sociometric test) and the process scale changes.

In order to test sub-hypothesis (2), correlations were run between change scores on all the psychometric measures of interpersonal attitudes. If the hypothesis is to be upheld, there should be a significant correlation between the F scale, as a measure of authoritarianism, and the WLP, Relationship Inventory, and sociometric test, as measures of interpersonal perceptions. This correlation should be a negative one, inasmuch as positive change on the F scale is indicated by a decrease in score.

The WLP was the only instrument to support the hypothesis of related changes. Five of the eight subscales showed significant change correlations with the F scale, and one other just missed significance at the .05 level. Thus positive change on those variables measured by the WLP does seem to be associated with a decrease in authoritarianism (see Figure V.21).

Neither the Relationship Inventory nor the sociometric test yeilded global support for the hypothesis, though some subscales of R.I. showed the expected correlations. Two of the four subscales for the Group ratings on the Relationship Inventory had significant negative change correlations with the F scale (Empathy and Unconditionality, with r.'s of -.473 and -.317, respectively). All but three of the 12 R.I. subscales in this comparison showed the expected negative correlation, but for most cases at a nonsignificant level; the three positive correlations were .137, .003, and .066, respectively.

Of the sociometric test changes, the only one which approached significance was the Comfort rating received by the subjects; here the correlation was +.216 (for a 2-tailed test an r. of .250 would be needed for significance). However, since the hypothesis being tested was clearly directional, even if the correlation value had exceeded the necessary level we would not be justified in calling it "significant."

As can be seen from the description of the testing procedures, the data gathered did not provide sufficient information to test hypothesis (3) of this set. While the gap is regrettable, it was felt that the disadvantages of carrying out the testing program as originally conceived (see Chapter II) provided sufficient reason for discarding this particular test.

E. These positive changes are long-lasting rather than temporary.

1. Students who have experienced positive change in any of the dimensions being considered will tend to maintain or increase their level of attainment in these dimensions over the time of the study.

   (To be completed.)

Source	SS	df	MS	F	p
Structure	.01	1	.01		
Congeniality	8140.61	1	8140.61	10.44	<.005
S x C	183.01	1	183.01		
Error	20270.55	36	563.07		
Trials	241.51	1	241.51	3.97	<.10
T x S	165.31	1	165.31	2.70	
T x C	25.31	1	25.31		
T x S x C	6.61	1	6.61		
Error	2189.75	36	60.83		

Figure V.1. Analysis of variance of F scale scores.

Congenial	Uncongenial
49.85	70.03

Pre	Post
61.67	58.20

	High Structure	Low Structure
Pre	60.25	63.10
Post	59.65	56.75

Figure V.2. F scale mean scores for congeniality, trials & structure x trials interaction.

		Most Liked								Least Liked								Group							
		R		E		C		U		R		E		C		U		R		E		C		U	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
HIGH STRUCTURE	Congenial																								
	Non-Congenial																								
LOW STRUCTURE	Congenial																								
	Non-Congenial																								

Figure V.3. Overall design for analysis of variance for Relationship Inventory.

Source	SS	df	MS	F	p
Structure	1283.44	1	1283.44		
Congeniality	1771.27	1	1771.27		
St x C	555.10	1	555.10		
Error	47444.11	36	1317.89		
M-L-G	37339.35	2	18669.68	152.68	<.005
Scales	62679.26	3	20893.09	170.86	<.005
Trials	382.54	1	382.54	3.13	<.10
M x Sc	10781.15	6	1796.86	14.69	<.005
M x T	62.67	2	31.33		
Sc x T	50.07	3	16.69		
M x Sc x T	712.45	6	118.74		
St x M	57.19	2	28.60		
St x Sc	451.50	3	152.50	1.25	
St x T	360.15	1	360.15	2.95	<.10
St x M x Sc	673.44	6	112.24		
St x M x T	347.64	2	173.82	1.42	
St x Sc x T	175.58	3	58.52		
St x M x Sc x T	150.71	6	25.12		
C x M	1784.58	2	892.29	7.29	<.005
C x Sc	1514.14	3	504.71	4.13	<.01
C x T	.50	1	.50		
C x M x Sc	233.09	6	38.85		
C x M x T	325.28	2	162.64	1.33	
C x Sc x T	78.39	3	26.13		
C x M x Sc x T	234.71	6	39.12		
St x C x M	476.50	2	238.25	1.95	
St x C x Sc	1664.09	3	554.70	4.54	<.01
St x C x T	601.67	1	601.67	4.92	<.05
Residual Error	103324.78	845	122.28		

Figure V.4. Analysis of variance for Relationship Inventory scores.

	R	E	C	U	
Most liked	63.91	40.34	48.31	32.90	46.37
Least liked	34.54	28.24	38.60	23.00	31.09
Group	48.41	34.04	43.44	28.29	38.54

Figure V.5. Mean scores for four subscales x M-L-G interaction on Relationship Inventory.

	M				L				G			
	R	E	C	U	R	E	C	U	R	E	C	U
Congenial	66.35	39.80	51.25	37.13	37.55	29.00	41.60	25.73	48.18	31.45	44.23	28.05
Non-congenial	61.48	40.88	45.38	28.68	31.53	27.48	35.60	20.25	48.65	36.63	42.64	28.52
	46.37				31.09				38.54			

	R	E	C	U	
Congenial	50.69	33.42	45.69	30.30	40.03
Non-congenial	47.22	35.00	41.21	25.82	37.31
	48.95	34.20	43.45	28.06	

Figure V.6. Mean scores for Congeniality x scales, and M-L-G x scales interactions on Relationship Inventory.

High Structure	Congenial	34.41	42.43
	Non-congenial	38.75	39.70
Low Structure	Congenial	40.43	38.83
	Non-congenial	34.56	36.23
		38.04	39.30

Figure V.7. Mean scores for Structure x Congeniality x Trial interaction on Relationship Inventory.

Source	SS	df	MS	F	p
Structure	11,559,634	1	11,559,634	5.53	<.05
Congeniality	11,270,259	1	11,270,259		
S x C	2,056,007	1	2,056,007		
Error	752,750,952	36	20,909,749		
Trials	20,682,729	1	20,682,729	3.80	<.05
S x T	7,782	1	7,782		
C x T	4,047	1	4,047		
S x C x T	4,485,466	1	4,485,466		
Error	196,050,996	36	5,445,861		

Figure V.8. Analysis of variance of Q sort z' transformations, self-ideal correlations compared pre- and post-treatment (data converted to integers by linear transformation).

Group	Mean score
High structure	.9635
Low structure	.7231
Pre	.7925
Post	.8942

Figure V.9. Mean scores for significant main effects in analysis of self-ideal correlations, pre- and post-treatment.

Source	SS	df	MS	F	p
Structure	19,426,176	1	19,426,176	2.04	<.10
Congeniality	21,974,465	1	21,974,465	2.31	<.10
S x C	794,011	1	794,011		
Error	343,130,154	36	9,531,393		
Self/ideal	9,823,815	1	9,823,815	3.31	<.05
S x S/I	24,376,320	1	24,376,320	8.21	<.01
C x S/I	1,238,526	1	1,238,526		
S x C x S/I	4,107,805	1	4,107,805		
Error	106,812,574	36	2,967,016		

Figure V.10. Analysis of variance of Q sort z' transformations self-self and ideal-ideal comparisons (data converted to integers by linear transformation).

Group	Mean score
High structure	.96358
Low structure	.86503
Congenial	.96672
Non-congenial	.86190
Self-self	.87926
Ideal-ideal	.94935

Figure V.11. Mean scores for main effects in analysis of self-self and ideal-ideal correlations.

	S-S	I-I
High structure	.98374	.94343
Low structure	.77478	.95527

Figure V.12. Mean scores for groups, structure x S/I interaction in analysis of self-self and ideal-ideal correlations.

Source					
Structure	52,509,062	1	52,509,062	10.14	<.01
Congeniality	15,740,463	1	15,750,463	3.04	<.05
S x C	2,860,183	1	2,860,183		
Error	186,399,593	36	5,177,766		
Trials	8,340,507	1	8,340,507	6.8	<.05
S x T	319,160	1	319,160		
C x T	722,190	1	722,190		
S x C x T	914,423	1	914,423		
Error	44,280,221	36	1,230,006		

Figure V.13. Analysis of variance of Q sort z' transformations, self-Expert pre- and post-treatment (data converted to integers by linear transformation).

Source					
Structure	49,850	1	49,850		
Congeniality	6,234,536	1	6,234,536	2.06	<.10
S x C	966,021	1	966,021		
Error	109,031,343	36	3,028,648		
Trials	47,580	1	47,580		
S x T	5,224,976	1	5,224,976	3.39	<.05
C x T	2,434,276	1	2,434,276		
S x C x T	3,128,800	1	3,128,800	2.03	<.10
Error	55,471,657	36	1,540,879		

Figure V.14. Analysis of variance of Q sort z' transformations, ideal-Expert pre- and post-treatment (data converted to integers by linear transformation).

	self-Expert	ideal-Expert
High structure	.71799	.81939
Low structure	.55595	.81440
Congenial	.59250	.84481
Non-congenial	.68134	.78900
Pre-treatment	.60468	.81933
Post-treatment	.66926	.81446

Figure V.15. Group means for main effects in analysis of self-Expert and ideal-Expert correlations, pre- and post-treatment.

		Pre	Post
High structure	Congenial	.84907	.86752
	Non-congenial	.8457	.71527
Low structure	Congenial	.81054	.85211
	Non-congenial	.77202	.82292

Figure V.16. Group means for interactions in analysis of ideal-Expert correlations, pre-and post-treatment.

PRE-TREATMENT PAIRS

High Regard	Low Regard
$\bar{X} = 4874.$	$\bar{X} = 4615.6$

pooled est. $\sigma^2 = 26141975.1$
est. $\sigma_{diff} = 2800.4$

$t = .09$ d.f. = 13

POST-TREATMENT PAIRS

High Regard	Low Regard
$\bar{X} = 5411.$	$\bar{X} = 3611.$

pooled est. $\sigma^2 = 7081935.45$
est. $\sigma_{diff} = 1303.5$

$t = 1.38$ d.f. = 17

Figure V.17. t tests for high and low mutual regard pairs, pre- and post-treatment, using z' scores from Q sort correlations (data converted to integers by linear transformation).

		Given						Received					
		Liking		Respect		Comfort		Liking		Respect		Comfort	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
High structure	Congenial												
	Non-congenial												
Low structure	Congenial												
	Non-congenial												

Figure V.18. Overall design for analysis of variance for sociometric test scores.

Source	SS	df	MS	F	p
Structure	5817.2	1	5817.2	12.7	<.01
Congeniality	12617.8	1	12617.8	27.5	<.01
St x C	596.3	1	596.3		
error	16519.1	36	458.9		
Given/Received	65.3	1	65.3		
Scales	584.6	2	292.3	3.3	<.10
Trials	408.9	1	408.9	4.6	<.05
G/R x Sc	8.4	2	4.2		
G/R x T	11.7	1	11.7		
Sc x T	127.6	2	63.8		
G/R x Sc X T	29.5	2	14.7		
St x G/R	51.3	1	51.3		
St x Sc	333.1	2	166.5	1.9	
St x T	312.0	1	312.0	3.5	<.10
St x G/R x Sc	104.2	2	52.1		
St / G/R x T	.35	1	.35		
St x Sc x T	747.1	2	373.5	4.2	<.05
St x G/R x Sc x T	.87	2	.43		
C x G/R	438.9	1	438.9	4.9	<.05
C x Sc	462.3	2	231.2	2.6	<.10
C x T	.02	1	.01		
C x G/R x Sc	22.05	2	11.0		
C x G/R x T	.47	1	.47		
C x Sc x T	108.9	2	54.5		
C x G/R x Sc x T	.35	2	.18		
St x C x G/R	8.3	1	8.3		
St x C x Sc	314.6	2	157.3		
St x C x T	1.5	1	1.5		
error	35800.8	403	88.8		

Table V.19. Analysis of variance for sociometric test scores.

	Congenial	Non-congenial	Overall
High structure	61.8	53.8	57.8
Low structure	71.0	58.5	64.8
Overall	66.4	56.2	

Figure V.20. Group means for nonrepeated-measures variables, analysis of variance for sociometric test scores.

	Pre	Post
High structure	59.5	56.1
Low structure	64.9	64.6
	62.2	60.4

Figure V.21. Group means for structure x trial interaction, analysis of variance of sociometric test scores.

	Pre			Post		
	L	R	C	L	R	C
High structure	59.2	61.0	58.5	56.6	53.7	58.0
Low structure	68.4	60.7	65.6	65.2	63.7	65.1

Figure V.22. Group means for structure x scale x trial interaction, analysis of variance of sociometric test scores.

	Given	Received
Congenial	67.0	65.8
Non-congenial	54.8	57.5

	L	R	C
Congenial	68.1	63.5	67.7
Non-congenial	56.6	56.0	55.9

Figure V.23. Interaction mean scores for congeniality with Given/Received and with scales; analysis of variance of sociometric test scores.

Adults

	21	34	43	46
15	.090	.329	.692	.126
21		-.066	.106	.173
34			.389	.071
43				.263
46				

Pity

	11	16	22	40
4	.183	.263	.435	.244
11		.140	.293	.188
16			.242	.228
22				.371
40				

Confidence

	17	24	36	42
2	.179	.196	.211	.177
17		.312	.126	.083
24			.159	.060
36				.241
42				

Giving

	12	27	31	39
10	.283	.350	.101	.255
12		.357	.349	.101
27			.225	.159
31				.263
39				

Liking

	14	19	28	47
7	.200	.377	.342	.180
14		.253	.317	.277
19			.291	.050
28				.136
47				

Values

	20	33	44	53
6	.244	.254	.272	.280
20		.380	.390	.445
33			.357	.555
44				.361
53				

Trust

	30	35	48	50
3	[illegible]	.478	.415	.309
30		.381	.405	.205
35			.310	.121
48				.214
50				

Basic nature

	23	32	37	51
8	.281	.375	.468	.185
23		.386	.295	.364
32			.288	.316
37				.122
51				

Figure V.24. Inter-item (within subscale) correlations for the WLP.

Source	SS	df	MS	F	p
Structure	46.1	2	23.1		
Congeniality	31.0	1	31.0		
S x C	74.5	2	37.2		
error	873.9	24	36.4		
Scales	727.7	7	104.0	14.4	<.01
Trials	8.0	1	8.0		
Sc x T	59.2	7	8.5		
St x Sc	180.2	14	12.9	1.8	<.05
St x T	10.6	2	5.3		
St x Sc x T	38.9	14	2.8		
C x Sc	37.2	7	5.3		
C x T	.8	1	.8		
C x Sc x T	14.5	7	2.1		
St x C x Sc	220.7	14	15.8	2.2	<.05
St x C x T	9.1	2	4.5		
St x C x Sc x T	22.1	14	1.6		
error	2583.3	360	7.2		

Figure V.25. Analysis of variance of WLP scores, using structure and congeniality as nonrepeated independent variables.

		A	P	G	C	I	V	T	B
Low structure	Congenial	13.4	16.6	15.0	11.5	12.3	16.4	16.2	15.0
	Non-congenial	11.6	14.3	15.8	12.0	14.1	14.6	15.8	13.6
Middle structure	Congenial	10.8	13.3	14.7	11.5	14.5	14.5	13.8	13.5
	Non-congenial	12.9	12.9	13.3	12.4	13.5	16.9	15.4	15.9
High structure	Congenial	15.2	12.0	14.6	11.6	13.8	15.0	14.2	14.3
	Non-congenial	13.7	14.9	16.4	11.4	14.6	15.5	17.1	17.3

Figure V.26. Group means for structure congeniality x scale interaction, analysis of variance of WLP scores.

Source	SS	df	MS	F	p
Schools	130.7	2	65.3		
Congeniality	7.0	1	7.0		
Sch x C	203.0	2	101.5	1.6	
error	3393.4	54	62.8		
Scales	955.4	7	136.5		
Trials	7.0	1	7.0		
Sc x T	60.3	7	8.6		
Sch x Sc	283.9	14	20.3	2.7	<.01
Sch x T	22.4	2	11.2		
Sch x Sc x T	18.7	14	1.3		
C x Sc	81.2	7	11.6		
C x T	.8	1	.8		
C x Sc x T	13.9	7	2.0		
Sch x C x Sc	162.0	14	11.6		
Sch x C x T	34.2	2	17.1	2.3	<.10
Sch x C x Sc x T	64.8	14	4.6		
error	6118.4	810	7.6		

Figure V.27. Analysis of variance of WLP scores, using school and congeniality as nonrepeated independent variables.

	A	P	G	C	L	V	T	B
Public	12.15	14.03	15.00	11.85	13.93	16.18	15.33	[illegible]
University	12.95	11.80	12.90	12.68	13.48	13.58	14.43	14.
Parochial	13.55	13.05	14.70	12.05	14.10	15.38	14.93	[illegible]

Figure V.28. Group means for schools x scales interaction analysis of variance of WLP scores.

Chapter VI

Miscellaneous Additional Analyses

A number of "peripheral" aspects of the research will be reported on in detail in the final report. For the purposes of the preliminary report, these peripheral studies will be grouped together on the following pages.

1. Normative study of the Q sort.

The normative data has yet to be analyzed. An analysis was carried out using the experimental data, and these results are reported here.

In order to determine whether the Q sort items which the Expert sort ranks at the same level tended to vary together within our population, correlations were run between each item pair across two Q sorts from each subject in the design. The expected clustering did not occur: the majority of the correlations among items given the same score on the Expert sort were not significant.

Again using the Expert sort as the criterion, a split-half reliability was computed across the subjects in the present design. Here the results were quite encouraging: with an n of 138, the split-half reliability was .856 (using the Spearman-Brown correction).

The results of these analyses, taken together, give us some insight into the way in which the Q sort works for this population. While the overall internal consistency is high, as evidenced by the high split-half r., individual item scores cannot be predicted. It seems that, within a very general framework of "this is like me, this isn't like me, this is about in the middle," there is great inter-item variability. Considering the varied--and varying--personality dynamics of this age group in general, such a finding is hardly surprising.

2. Normative study of the Relationship Inventory.

To be completed.

3. Comparison of high school and college students.

As one of the additional investigations which the "serendipity" features of the present design allow, it was planned to study the similarities and differences between the present population of eleventh-graders and the college freshman population of O.E. project #1417. Several selection features in the two samples may contribute to whatever differences occur in the data (volunteer vs. non-volunteer subjects; homogeneity of geographical background; college interests; male/female ratio), and these should be taken into account in any interpretation of the differences which occur.

In view of the many factors which might have made for systematic differences between the two samples, the actual number of statistically significant differences is surprisingly small. The F scale revealed no significant dffferences at all between the two populations. One difference shows up on the Relationship Inventory, within the category of "person most liked in group" (the third category did not provide a valid comparison, inasmuch as it was not defined in the same way for the two populations: the high school group rated "group as a whole" while the college students rated the group leader). Here both boys and girls in the high school group rated their favorite person higher than did those in the college group. This might be interpreted in terms of higher intensity of teenage relationships, with less qualification of regard, and/or the fact that many of the college groups were all female rather than heterogeneous with regard to sex. Any such interpretation, however, is subject to doubt because of the multiple t tests run across the R.I. scores. A single significant t value out of 32 tests can hardly be regarded as upholding a theoretical notion of this sort.

The MMPI showed more consistent results. On the Hs, Pd, Pt, Sc, and--to a lesser degree--Ma, the high school students were significantly higher than the college group. Assuming that the MMPI is a valid instrument for both age groups, the notion of differing personality dynamics between the groups is supported. The high school and the college students would seem to have displayed some qualitative differences here. (A later section will deal with the application of an especially derived subscale of the MMPI which has as its goal the prediciton of positive movement in counseling.)

By far the most striking differences between the groups occurred on the Q sort. Here the high school students clearly showed higher correlations between self and Expert sorts than did the college students. We cannot be sure whether the college students' relatively poor performance was attributable, at least in part, to test-taking attitudes rather than to more basic personality differences.

In summary, then, three of the four measures checked show, to varying degrees, differences between the high school and the college populations. These differences were not consistently in favor of one group or the other. The Relationship Inventory and the Q sort both showed the high school students in a more favorable light: as perceiving both selves and peers more positively than did the college students. The MMPI, on the other hand, showed a tendency for the high school subjects' scores to be significantly higher; this generally indicates less healthy personality dynamics.

A number of factors may have contributed to these contradictory effects: The idealism of the younger group as opposed to the psuedo-sophistication so often donned by the college freshman; a difference in test-taking attitudes; different reactions to the test materials themselves. Last, but clearly not least, it must be remembered that the high school students were volunteer subjects, while the college population subjects were required to participate as a part of one of their regular academic courses.

It would be pleasant to be able to conclude from the above analyses that the high school students tended to be more open and acceptant toward peers and selves, yet at base less mature and stable in personality make up. While the results may point slightly in this direction, such a generalization must be regarded as a hypothesis; in no sense of the word may it be inferred that the hypothesis has been supported or "proved" by the data.

4. Utilization of the MMPI as a predictor of "success" in counseling.

To be completed.

5. Development of content-Oriented tape rating scales.

During the course of its work in the area of psychotherapy with schizo-

During the course of its work in the area of psychotherapy with schizophrenics, the Psychotherapy Research Group of the Wisconsin Psychiatric Institute has developed a number of scales designed to measure the progress, or process, of psychotherapeutic change. Originally conceptualized by Carl Rogers and colleagues at the University of Chicago, the Process Scale began as a single, wholistic measure of an individual's position on a theoretical spectrum of psychological well-being. Gradually it became apparent that the wholistic approach did not provide an adequate measure, and the Process Scale was broken into "strands", corresponding roughly to those aspects of therapy behavior considered most relevant to psychological "improvement." While considerable addition to and refinement of these strands is still going on, three in particular have emerged as useful in the area of counseling and have been rather extensively used in the client-centered framework. These have also been used in the analysis of our present study.

The Process Scales, however, have one serious drawback as regards their applicability to the type of data being gathered in this study. They are designed for the analysis of individual counselee behavior in the one-to-one counseling or therapy situation. They contain little or no relational referent material, and no items or descriptions which take in to account the degree and manner in which the individual is interacting with other members of the group--a factor quite important in both the progress and the assessment of group counseling. For this reason, it was necessary to devise new scales directed primarily toward evaluating this group interaction aspect. Such scales would be rather specifically oriented toward interpersonal interactions, and would be used in conjunction with the other, individually oriented scales.

Two scales evolved from our thinking and discussions about those kinds of interaction which we felt were helpful or useful or "therapeutic" in a counseling group. Scale PAR represents an attempt to measure the amount of participation in group discussion which a given individual is doing. At the lower stages it is concerned solely with whether or not the individual is making himself heard as part of the group; moving upwards there is an increasing attention to the kind of participation, with the expectation that as the group and its members experience successful counseling, the individual will more and more tend to express things which are personally meaningful to him, and will exhibit more and more flexibility in the manner of their expression and discussion.

Scale COM concerns itself with the other side of the coin: listening, rather than talking, to the group. At stage 1 the individual cares only about what he has to say and listens minimally to the others. Progressing up the scale, there is more and more attention to what others are saying, and an increasingly empathic response to verbalizations from fellow group members.

A number of studies have been made of the optimal length of segment for accurate and efficient rating. A section in the Final Report will discuss in detail the process of segment-selection used for the individually oriented scales. For the group interaction scales, however, it was felt that a considerable longer segment was needed to allow the rater a chance to "feel" what the group was doing during a given session. Also, the temper of the group could change radically over a relatively short period of time: from boistrous kidding to serious talk, or vice-versa. Finally, it was suspected that the group as a whole could be rated on the scales, and that this rating might differ considerably from any average or weighted combination of the individual ratings. For these reasons, it was decided to allow the raters to listen to the whole hour tape of the interview to be tated, and keep a running tally of peak ratings for each of the six participants and for the whole group. The ratings were made for all groups at five selected data points.

Inter-rater reliabilities were figured for possible pairs selected from the four raters. Reliabilities for individual ratings were quite respectable for both scales, averaging around .6 over better than 150 observations. In rating the groups as a whole, only scale PAR showed acceptable inter-rater reliability (averaging around .7); reliabilities on COM were little better than chance.

SCALE PAR

This is a scale to help you to rate the kind of participation that a person gives in the interaction of the group. Familiarize yourself thoroughly with the scale stages described below, and try to apply them to the verbal behavior of the individuals in the group in as objective a manner as possible.

STAGE 1.
The person doesn't talk at all, or else gives minimal responses to direct questions.

STAGE 2.
All talk is social. Giggles, standard "parties-dating-movie" talk. Personal anecdotes may be told as jokes. The speaker seems to be taking the role of an actor or entertainer.

STAGE 3.
Speech is still very superficial, but personal anecdotes are related now. A personal reaction to an event or to what someone else is saying may occur (but this is very casual, usually in joking tones).

STAGE 4.
Person talks relatively seriously about events, experiences. The comments are not personalized. He may repeat an opinion but he does not claim it as his own.

STAGE 5.
Opinions, attitudes, and conclusions are given in an intellectualized fashion. Objective facts may be used to support a statement made. There is little or no use of personal experiences to help the person explain what he means.

STAGE 6.
Both facts and feelings are used to support opinions. Personal experiences and reactions are used to support opinions and to explain what the speaker means.

STAGE 7.
(This is essentially an extension of Stage 6)
The person may indicate a willingness to have his own viewpoint altered or to accomodate a new and different point of view. Even when he is quite certain of his own stand (and if so, he will be able to support it clearly and logically) he is respectful of others' points of view. What he is saying is important and serious to him, and he is trying hard to make himself understood.

SCALE COM

This is a scale designed to help you rate the kind of communication which is going on in the group. In order to rate a given individual accurately, you will have to take into account whatever it is he is replying to. Listen carefully, and try to be as objective as possible in your ratings.

STAGE 1.

Speaker totally ignores whatever point(s) others in the group may be trying to make. He changes the subject, interrupts with a totally new train of thought, etc.

STAGE 2.

The speaker uses another person's remark(s) as a stepping-stone for his own, but he is clearly not interested in what the other is saying. He listens primarily in order to find a spot where he can break in and tell what he is interested in.

STAGE 3.

Either the speaker tries to "beat down" the other(s) who are talking, and shows interest in them only insofar as to get them to admit that he (the speaker) is right

or the speaker participates in "semi-hysterical" interactions, with much giggling, yelling, etc.

STAGE 4.

The speaker seems to "side" with one other person and can understand and rephrase his ideas. But this occurs with only one other person, and only when the ideas expressed agree substantially with his own. He is still very intolerant of any opposing ideas and tends to ignore them or argue very defensively when he is forced to listen to them.

STAGE 5.

The speaker is trying to understand what the other people mean, but he still is not aware of how they feel in an empathic sense. He isn't really vitally interested in their opinions, and his listening is limited in terms of his attention span. When he responds it is at a very intellectual level.

STAGE 6.

The speaker is actively involved in understanding the feelings and reactions of the other members of the group. He may request more information and/or indicate both awareness of and respect for what the others are saying.

Chapter VII

Conclusions

It is at this point incumbent upon the authors to attempt to pull together from the findings reported in the previous pages some sort of unified picture. In a situation involving as many variables as does the present research, such a unified picture must, of necessity, sacrifice some degree of precision in exchange for conciseness and generality. We shall, in the following pages, comment in as specific and accurate a way as possible upon the research findings; nevertheless, it should be recognized that this chapter serves primarily as a point of departure for hypothesizing and speculation rather than detailed analyses of results. For the latter type of discussion, the reader is referred to Chapter V of this report.

In this chapter we shall consider in turn the independent variables which have been utilized throughout the analyses. The discussion will center first upon the results of the psychometric data, and then upon the analyses of the group meetings themselves.

The factor of congeniality has consistently been of major concern in the research plan. While it seems clear that group congeniality, as such, has been confounded with individual popularity and peer esteem, this confounding is not necessarily a drawback. In the first place, it may be that naturally congenial groups tend to sort themselves out more often among the highly "popular" teenagers than otherwise. If so, any attempt to separate congeniality from "popularity" in the experimental situation might well result in the creation of an artifiality of grouping which would yield data relatively ungeneralizable to the real world of the student school population.

Secondly, assuming that at the level of natural groupings congenial groups do form as frequently among the less popular students as among the more popular, the situation found in the present research still reflects the kind of groups the counselor can expect among volunteers for group meetings. That is, it is much more likely that the more popular, self-confident students would formalize their group structure for the purpose of "group discussion" than that the less popular students would do so. The latter would more probably hesitate to coalesce the group or to verbalize among themselves the idea that a group did, in fact, exist--the stumbling-block being each individual's lack of security in his perception of the esteem of the others.

The outstanding exception to this logic is, of course, the delinquent or border-delinquent group. Here the solidarity of the "gang" is manifest and, indeed, stands as a bulwark against the negative sanctions of the rest of the rest of the adolescent society. However, this type of group member is quite unlikely to present himself on a voluntary basis for the kind of group activity with which our research is concerned. It must be remembered that the whole framework and context of the present research is that of volunteer subjects. No generalization to other than volunteer participation is implied, or should be inferred.

The psychometric data leave little doubt as to the existence of significant differences between the congenial and the non-congenial groups. Repeatedly, the congenial group subjects' scores are more positive, reflecting generally more favorable attitudes and perspectives of themselves and the world around them. The conclusion that young people of this sub-culture tend to select out among themselves those individuals whose attitudes and values are--in the opinion of the adult world--most healthy and mature seems inescapable. Data from the F scale, the Q sort, the sociometric test, and the teacher ratings all support such a conclusion.

An outstanding exception to this trend is the relative standing of congenial and non-congenial groups on the self-sort of the Q sort. Here alone the non-congenial group data show up more positively than the congenial:

the former correlate more highly with the Expert sort, indicating a view of self closer to the Expert's projection of the "best possible self." A question may be raised, of course, as to the honesty of this self-report. The authors, however, are more inclined to view the relationship as reflecting a higher degree of introspection and self-criticism, and possibly a higher goal or ideal standard among the congenial group students. The fact that the correlation of ideal self with Expert was higher among the congenial subjects tends to support this observation.

In summary, then, there emerges a picture of the type of student who tends to be "chosen out" by his peers. He is held in esteem not only by peers, but by his teachers as well. His attitudes are relatively nonauthoritarian, and he tends to be accepting of and comfortable with his peers. His standards for his own personal and personality development are high, and he is insightful and/or critical in analyzing his progress toward meeting those standards.

The treatment variable, group meetings, can be evaluated in terms of its overall effect through an examination of the pre- and post-treatment F ratios in the several analyses of variance. All of the psychometric tests administered pre- and post-treatment, with the exception of the WLP, show significant differences between the two testing periods. It is clear that changes did occur.

Author's note: this section is incomplete; it will be completed upon finishing the analyses of the control and follow-up data.

In general, changes over time were of a positive nature. Scores on the F scale and the Relationship Inventory indicated an overall decrease in authoritarianism and an increase in perception of "therapeutic" qualities within the group. In contrast, however, the more general ratings of fellow group members, as seen in the sociometric test, went down over the course of the group sessions. This seeming contradiction occurred primarily among high-structure group members, and will be discussed more fully in the next section.

Information yielded by the various Q sort analyses indicates general support to the hypotheses of positive changes over time, although different structure level groups manifested these changes in different ways. The overall nature of the changes was in the direction of greater correspondence between both self and ideal sorts and the Expert sort.

The trend of over-time changes, as reflected in the psychometric data, can be summarized: the students seemed to become more accepting and tolerant of others and of themselves, and to shift their internal standards of personal development in the direction of a morae mature sort of goal or image.

While many of the trends summarized in this section failed to reach the level of statistical significance generally considered acceptable, the total picture is relatively homogeneous. It is this homogeneity which, in the opinion of the authors, justifies the reporting of results in this fashion. We have emphasized, in previous chapters, the hypothesis-generating nature of the research. It would be quite unwarranted to regard these results as "proof" that group counseling does create positive changes among an adolescent volunteer population. The evidence suggests that such an effect does, indeed, occur; and in our opinion is ample justification for further study. It is this further study, however, rather than the results on these pages, which may yield the confirmation of our hypotheses.

The effect of structure level upon progress in group counseling has been perhaps the main focus of the research. It is at the point of examining this relationship, unfortunately, that the orderliness of the results begins to break down. In view of the complexity of the variables called into play as an attempt is made to define and/or control the structuring of group discussion, such apparent lack of order is perhaps not surprising. "Structure" means many things to many people, and each of the individuals involved in the study must undoubtedly have developed a unique perception of and reaction to the structure level under which he operated.

The most surprising results with regard to structure level were those which indicated overall differences between the different structure groups. As was mentioned earlier, the existence of an overall main effect in what was essentially a treatment variable is difficult to understand. In the analyses of Q sort data, the high structure groups tended to produce higher correlations with the Expert sort. In the sociometric test, the low structure groups reflected a generally more positive perception of fellow group participants. The latter effect may be due to the immediately apparent nature of the structuring and the effect which the structure levels may have had upon the social nature of the newly forming group. The Q sort result, on the other hand, remains something of a mystery; any attempt to explain or dissect it would be little more than speculation. In a later paragraph we shall consider the Q sort analyses more specifically.

Considering the relation of structure to changes over time, the general tendency seemed to be that of more positive changes on the part of low structure groups. This trend appeared in the F scale data, in the correlation of ideal Q sort with Expert, and in the sociometric test (in the latter, it was not so much positive changes of the low structure groups as decrease in score among high structure groups which contributed to the interaction). In contrast, the high structure groups improved their scores more than did the low structure groups on the Relationship Inventory, and the correlation of pre- with post-treatment self Q sorts indicates significantly more movement (positive or negative direction not specified) on the part of low structure group members. In view of these contradictions it would seem highly premature to attempt to pull out some consensus of trend in the effect of structure alone. The two specific measures of social perception, the Relationship Inventory and the sociometric test, show results in direct contradiction to each other; the WLP, a measure of social attitudes, picks up nothing.

By far the most complex relational picture of the overall results was presented by the Q sort. While these results were discussed in detail in Chapter V, it may be of value to re-summarize here. The emphasis on the Q sort springs in part from the fact that it has allowed of a more detailed analysis than any other single psychometric instrument, and partly from the authors' inclination to regard it as particularly sensitive under the present research conditions. Considering only changes in the subjects' sorts, without reference to the Expert standards, there tended to be more movement or change over time among the non-congenial group members. The low structure groups, across the board, also tended to show a greater degree of movement. Yet the only Q sort measure which showed any interaction effect between structure and congeniality over time was the correlation between Expert and ideal self: the high structure non-congenial subjects showed negative change here while the other groups changed positively.

It is clearly not justifiable at this point to point to our results as clear support of any hypothesis. Nevertheless, some tentative outlines may be present. On no measure did the low structure congenial groups do well. The high structure non-congenial groups not only failed to surpass any of the others, but actually showed negative change in one instance. Yet both the high structure congenial and the low structure non-congenial seem to have made fairly consistent progress. If any hypothesis emerges from the welter of data, it must be based on these observations: high structure tended to facilitate the congenial groups, while the reverse was true for the non-congenial groups.

It is our supposition that the non-congenial group members may have utilized the group setting to strengthen and enhance their social skills and their perceptions of themselves as social beings. The low structure setting would have been optimal for such activity. The congenial group members on the other hand, had little need for this sort of enhancement; their social skills and self-perceptions were objectively quite adequate. Here the lack of structure allowed the group meetings to digenerate into gossip-and-giggle sessions; only under high-structure conditions were the subjects led into the kind of group activity which could bring about positive changes.

Appendix

Computer Techniques

Contained in this appendix are programs and program descriptions which have been found useful in handling many of the standard statistical problems which arise in educational research.

The Programs are written in FORTRAN 60, and should, for the most part, be also suitable for a FORTRAN 63 or a SCOPE compiler.

Program descriptions cannot, of course, include specific instructions for setting up a deck to run under a given compiler system. The user should familiarize himself with the restrictions and requirements of the particular computer system with which he is working.

While these programs have been checked out on either the CDC 1604 or the CDC 3600, no guarantee of their accuracy is given or implied by the author.

Program SPLHF

Q sort technique is adaptable to a wide variety of clinical and experimental situations. Simply by substituting new items, the focus of the instrument can be directed in any manner of use to the tester. However, such changes may have unknown effects on the reliability and validity of the instrument.

A single Q sort, in and of itself, does not yield a useable "score." The technique is most often used in two ways: comparing, by means of correlation, pairs of sorts between testees or between different administrations of the sort to the same testee; and comparing a testee's sort with some external standard. This latter method is of particular use in that it yields scores which can then be compared among testees in much the same way as scores on other, more traditional tests.

Program SPLHF was constructed in order to calculate split-half reliability measured on an 80-item, forced-normal Q sort. The reliability r. calculated by SPLHF is based on within-subject comparisons of correlations between subject and "expert" sorts.

While the program as written can be used only for an 80-item sort, only a few statements need be changed in order to make it apply to any even-numbered item sort.

Directions for use:

1. "Expert" sort. The sort which is to be used as a standard for comparison of the testee sorts is read in first. Since there are 80 items, and each one receives a score of 0 to 8, the entire expert sort can be punched on a single card, one item per column. This is most readily done by transferring the expert sort from the report form to the scoring form, and punching directly from the latter.
   The punched expert sort card is placed directly following the program deck.

2. Data. Data is of exactly the same form as for program QGRPS (see above). The data follows the expert sort immediately, and up to 200 sorts may be used. Following the last data card must be:
   - a) end signal card -- contains -1 in cols. 9 and 10
   - b) two blank cards

Note: This program utilizes an extra card following the final END card. This card contains the "EXPERT" sort, and is treated by the program as a data card; it should be kept as a part of the program deck. In the program given here, the "EXPERT" sort is that devised by Dymond for the Rockerfeller 80-item sort.*

---

*Dymond, Rosalind and Rogers, C. R. Psychotherapy and Personality Change, University of Chicago Press, 1954.

```
      PROGRAM SPLHF
      DIMENSION EXP(80)
      DIMENSION ID(10), SCORE(80), RAW(80), A(200), B(200)
      K =0.
      READ 12, EXP
12    FORMAT(80F1.0)
      DO 122 I=1,80
122   SCORE(I)=4.
31    READ 1, RAW
1     FORMAT (2(8X,35F2.0/),8X,10F2.0)
      IF(RAW(1))70,10,10
10    DO 2 I=1,3
      JJ=RAW(I)
2     SCORE(JJ)=0.
      DO 3 I=4,9
      JJ=RAW(I)
3     SCORE(JJ)=1.
      DO 4 I=10,18
      JJ=RAW(I)
4     SCORE(JJ)=2.
      DO 5 I=19,31
      JJ=RAW(I)
5     SCORE(JJ)=3.
      DO 6 I=32,49
      JJ=RAW(I)
6     SCORE(JJ)=4.
      DO 7 I=50,62
      JJ=RAW(I)
7     SCORE(JJ)=5.
      DO 8 I=63,71
      JJ=RAW(I)
8     SCORE(JJ)=6.
      DO 9 I=72,77
      JJ=RAW(I)
9     SCORE(JJ)=7.
      DO 11 I=78,80
      JJ=RAW(I)
11    SCORE(JJ)=8.
      XA=0.
      XB=0.
      XXA=0.
      XXB=0.
      YA=0.
      YB=0.
      YYA=0.
      YYB=0.
      XYA=0.
      XYB=0.
      DO 21 I=1,79,2
      YA=YA+EXP(I)
21    YYA=YYA+EXP(I)*EXP(I)
      DO 22 I=2,80,2
      YB=YB+EXP(I)
22    YYB=YYB+EXP(I)*EXP(I)
```

```
      DO 20 I=1,79,2
      XA=XA+SCORE(I)
      XXA=XXA+SCORE(I)*SCORE(I)
20    XYA=XYA+SCORE(I)*EXP(I)
      DO 30 I=2,80,2
      XB=XB+SCORE(I)
      XXB=XXB+SCORE(I)*SCORE(I)
30    XYB=XYB+SCORE(I)*EXP(I)
      K=K+1
      TA=40.*XYA-(XA*YA)
      BXA=40.*XXA-XA*XA
      BYA=40.*YYA-YA*YA
      BA=SQRTF(BXA*BYA)
      A(K)=TA/BA
      TB=40.*XYB-(XB*YB)
      BXB=40.*XXB-XB*XB
      BYB=40.*YYB-YB*YB
      BB=SQRTF(BXB*BYB)
      B(K)=TB/BB
      PRINT 112, K, A(K), B(K)
112   FORMAT(9HOSUBJECT ,I4,12H R VALUES = ,2F10.6)
      GO TO 31
70    DO 40 I=1,K
      A(I)=(.5*(LOGF(1.+A(I))-LOGF(1.-A(I))))*100.
40    B(I)=(.5*(LOGF(1.+B(I))-LOGF(1.-B(I))))*100.
      EN=K
      X=0.
      XX=0.
      Y=0.
      YY=0.
      XY=0.
      DO 41 I=1,K
      X=X+A(I)
      XX=XX+A(I)*A(I)
      Y=Y+B(I)
      YY=YY+B(I)*B(I)
41    XY=XY+A(I)*B(I)
      TOP=EN*XY-X*Y
      BOTX=EN*XX-X*X
      BOTY=EN*YY-Y*Y
      BOT=SQRTF(BOTX*BOTY)
      R=TOP/BOT
      PRINT 100, R, EN
100   FORMAT(16HOSPLIT HALF R = ,F16.5/14H N OF GROUP = ,F10.0)
      STOP
      END
      END
```

Program TMATCH

The computational procedures for t tests on correlated and on noncorrelated data vary considerably. Program TMATCH was designed to find t values for problems involving correlated data.

Identification card:

The first card following the program deck is the I.D. card. The first 16 columns on this card are printed on the output sheet exactly as punched. columns 17-20 contain the number of pairs of observations to be considered in the problem.

Format card:

Next is a standard FORTRAN format statement, with the word "format" omitted and first and last parentheses anywhere on the card.

Print format card:

The print format gives instructions for printing out an echo check of the data. It is punched in the same way as the format card.

If no echo check is desired, this card, plus the 7th and the 11th cards of the program deck, may be omitted.

Data:

Data is punched according to the format specifications. All of the X variables must be presented first, followed by all the Y variables, and no separator card between the X and Y variables. Ordering within X and within Y must correspond.

If more than one problem is to be worked, all control and format cards must be repeated, in the same order as for the first. Formats and n's need not be the same for subsequent problems.

Output:

In addition to the data echo check, the program yields values of t, N, and σ_{diff}.

Limits:

No more than 5000 pairs of observations are allowed for any one problem.

```
      PROGRAM TMATCH
      DIMENSION FMT(10), FMTP(10), X(5000), Y(5000), IDENT(2)
10    READ 2, IDENT, N
2     FORMAT (2A8, I4)
      IF(N) 41,41,40
40    READ 1, FMT
      READ 1, FMTP
1     FORMAT (10A8)
      READ FMT,(X(I),I=1,N)
      READ FMT,(Y(I),I=1,N)
      PRINT FMTP,(X(I),I=1,N),(Y(I),I=1,N)
      AN=N
      SUMX=0.
      SUMY=0.
      SUMD=0.
      SUMDSQ=0.
      DO 3 I=1,N
      SUMX=SUMX+X(I)
      SUMY=SUMY+Y(I)
3     SUMD=SUMD+ABSF(X(I)-Y(I))
      AVGD = SUMD/AN
      DO 20 I=1,N
20    SUMDSQ=SUMDSQ+(AVGD-(X(I)-Y(I)))**2
      SIGMA=SQRTF(SUMDSQ/(AN*(AN-1.)))
      AVGX=SUMX/AN
      AVGY=SUMY/AN
      T=ABSF(AVGX-AVGY)/SIGMA
      PRINT 11, IDENT, T
11    FORMAT(7HOT FOR ,2A8,3H = ,F6.3)
      PRINT 12, SIGMA, N
12    FORMAT(9H SIGMA = ,F10.3/5H N = ,I4)
      GO TO 10
41    STOP
      END
      END
```

Program BARTLETT

It is sometimes desirable to run a Bartlett's test to determine the degree of heterogeneity of variance in a set of data. This procedure is rather lengthy if done on a desk calculator; the program will handle any design using up to 400 data cards, with up to 20 items per card.

Control card:

The control card contains the following information.

cols. 1-4	number of cards to be read
cols. 5-8	number of cards per subject
cols. 9-12	number of cells in design
cols. 13-16	number of problems to be run
cols. 17-20	number of items per card
cols. 21-24	number of items per cell.

Format card:

This is a standard Fortran format statement, omitting the word "format," with first and last parentheses anywhere on the card.

Data:

Data should be punched in accordance with the Format card, and ordered in groups with each group consisting of the observations within one cell. There must be an equal number of observations in each cell. No separation cards are needed between observations, between cells, or between problems. When the cards for the first problem are exhausted, the program will either go on to the cards for the next problem or terminate, depending on the information given in columns 13-16 of the Control card.

No final termination or "end of job" card is required. The card containing the last observation in the last cell is the final card of the deck.

```
      PROGRAM BARTLET
C     SEE EDWARDS PAGE 196 FOR DESCRIPTION OF BARTLETTS TEST
C     MM IS NUMBER OF CARDS TO BE READ
C     NN IS NUMBER OF CARDS PER PERSON
C     LL IS NUMBER OF GROUPS
C     KK IS NUMBER OF PROBLEMS TO BE RUN
C     II IS NUMBER OF ITEMS(DATA POINTS) PER CARD
C     N1 IS NUMBER OF ITEMS PER GROUP
      DIMENSION A(420,22),B(12,5),C(12),D(12),E(12)
   10 FORMAT(11X,22F3.0)
    9 FORMAT(6I3)
      READ 9, MM,NN,LL,KK,II,N1
      READ 10,((A(I,J),J=1,II),I=1,MM)
      RL=LL
      N2=N1-1
      RN1=N1
      RN2=N2
      COR=1.+(RL+1.)/(3.*RL*RN2)
      DO 100 I=1,KK
      IF(I-II) 25,25,11
   11 IF(I-2*II) 26,26,12
   12 IF(I-2*II) 27,27,13
   13 IF(I-4*II) 28,28,14
   14 IF(I-5*II) 29,29,15
   15 IF(I-6*II) 30,30,16
   16 IF(I-7*II) 31,31,17
   17 IF(I-8*II) 32,32,18
   18 IF(I-9*II) 33,33,34
   25 N=1
      JJ=I
      GO TO 40
   26 N=2
      JJ=I-II
      GO TO 40
   27 N=3
      JJ=I-2*II
      GO TO 40
   28 N=4
      JJ=I-3*II
      GO TO 40
   29 N=5
      JJ=I-4*II
      GO TO 40
   30 N=6
      JJ=I-5*II
      GO TO 40
   31 N=7
      JJ=I-6*II
      GO TO 40
   32 N=8
      JJ=I-7*II
      GO TO 40
   33 N=9
      JJ=I-8*II
      GO TO 40
34    N=10
      JJ=I-9*II
40    DO 41 J=1,LL
      DO 41 K=1,N1
      III=N+NN*(N1*(J-1)+K-1)
41    B(J,K)=A(III,JJ)
      DO 51 J=1,LL
      S=0.
      T=0.
      DO 50 K=1,N1
      R=B(J,K)
      S=S+R*R
50    T=T+R
      C(J)=S-T*T/RN1
      D(J)=C(J)/RN2
      IF(D(J)) 110,110,51
110    D(J)=10.
51    E(J)=LOG10F(D(J))
      U=0.
      V=0.
      DO 52 J=1,LL
      U=U+D(J)
52    V=V+E(J)
      DIF=RL*LOG10F(U/RL)-V
      XX=2.302585*RN2*DIF
      ANS=XX/COR
8     FORMAT(I3,3X,I3,3X,I3,3X,3F18.6)
      DO 71 K=1,LL
71    PRINT 8,(K,N1,N2,C(K),D(K),E(K))
7     FORMAT(/20HANSWER FOR PROBLEM  ,I3,6H  IS  F10.4///)
100    PRINT 7,I,ANS
      STOP 9
      END
      END
```

Program UNEQ

While standard analysis of variance programs for a number of designs are available at most university computer centers, these programs usually require equal n's within cells. Inasmuch as random selection to equalize n's is often necessary in psychological and educational research, and often causes serious reduction of the overall n, this restriction may cause a relatively large increase in errors of inference.

With this in mind, it was decided to program the unweighted means analysis of variance method described by Winer*. This model assumes the differences in n to be caused by random fluctuations or by factors not related to the variables incorporated in the design. The program as it now stands will handle any p by q replicated design, excluding repeated measures. While Winer has extended the computational procedure to more levels, and to the repeated measure situation, our program does not attain that degree of generality.

Identification:
The first card following the program deck is the I.D. card. This may contain any letters, numbers, or symbols, and is printed out in full at the beginning of the output.

Control card:
The control card follows the I.D. card and gives the number of rows and columns in the design.

cols. 1-4 number of rows in design
cols. 5-8 number of columns in design

Format card:
Following the control card is a standard FORTRAN format statement, omitting the word "format" and with first and last parentheses anywhere on the card.

Data:
Data is punched according to the format statement. Only one observation is allowed per card. Cards are grouped by cells, with a card with -1 in the place designated for data by the format card following each cell. Cells, in turn, are ordered as follows (where R indicates rows and C indicates columns):

$$R_1C_1,\ R_1C_2, \ldots R_1C_n,\ R_2C_1,\ R_2C_2 \ldots R_nC_1 \ldots R_nC_n$$

Example:
Consider the following set of data.

	C_1	C_2	C_3
R_1	25	21	15
	20	19	20
			17
R_2	22	25	32
	23	26	33
	24		24

The data is punched, one number per card, in columns 3 and 4. Program UNEQ might require, following the program deck, these cards ("*" represents a blank space):

```
I.D. card       | 2 x 3 UNEQUAL  N  ANOVA FOR DEMO ANALYSIS
Control card    | *** 2 *** 3
Format card     | (8X, F3.0)
Data cards      | ** 25
                | ** 20
                | ** -1
                | ** 21
                | ** 19
                | ** -1
                | ** 15
                | ** 20
                | ** 17
                | ** -1
                | ** 22
                | ** 23
                | ** 24
                | ** -1
                | ** 25
                | ** 26
                | ** -1
                | ** 32
                | ** 33
                | ** 34
                | ** -1
```

Limits:

1) Only one observation per card.
2) Negative data is not allowed.
3) No more than 10 rows and 10 columns per problem.

Output:

In addition to sums of squares, degrees of freedom, and mean squares for rows, columns, and interaction, the program also yields the harmonic mean and the 5 computation terms listed by Winer. Row means and column means are printed out. As a diagnostic aid, a summing check matrix containing ΣX, ΣX^2, n, and sum of squares for each cell is also given.

*Winer, B. J. *Statistical Principles in Experimental Design*. McGraw-Hill Book Co., Inc., New York, 1962.

```
      PROGRAM UNEQ
      DIMENSION X(11,11),C(100,4),FMT(10)
      DIMENSION ROM(10), COM(10), ID(10)
      READ 2, ID
      READ 1, RO, CO
1     FORMAT(2F4.0)
      READ 2, FMT
2     FORMAT(10A8)
      PRINT 347, ID, FMT
347   FORMAT(1X,10A8)
      DFE=0.
      SSI=0.
      SSC=0.
      SSR=0.
      SSWC=0.
      S=0.
      SS=0.
      N=0.
      SD=0.
      SSD=0.
      EN=0.
      DO 200 I=1,11
      DO 200 J=1,11
200   X(I,J)=0.
24     I=0
      S=0.
      SS=0.
25    READ FMT, D
      IF(D) 16,10,10
10    S=S+D
      SS=SS+D*D
      I=I+1
      GO TO 25
16    N=N+1
      C(N,1)=S
      C(N,2)=SS
      C(N,3)=I
      C(N,4)=SS-(S*S)/C(N,3)
      PRINT 701, N, (C(N,I),I=1,4)
      K=RO*CO
      IF(N-K) 24,30,30
30    DEN=0.
      DO 35 I=1,K
35    DEN=DEN+1./C(I,3)
      HM=(RO*CO)/DEN
      PRINT 702,HM
      CK=0.
      II=RO
      JJ=CO
      L=0
```

```
         DO 400 I=1,II
         DO 400 J=1,JJ
         L=L+1
400      X(I,J)=C(L,1)/C(L,3)
         DO 411 I=1,10
         DO 411 J=1,10
         X(I,11)=X(I,11)+X(I,J)
411      X(11,I)=X(11,I)+X(J,I)
         DO 412 I=1,10
         X(11,11)=X(11,11)+X(11,I)
412      CK=CK+X(I,11)
         CK=ABSF(CK-X(11,11))
         IF(CK-1.) 20,20,21
21       PRINT 99
99       FORMAT(26H ERROR IN X MATRIX ROUTINE     )
         PRINT 98,((X(I,J),J=1,11),I=1,11)
98       FORMAT (5X,11F10.3)
         STOP
20       PRINT 703
         PRINT 704,((X(I,J),J=1,11),I=1,11)
         K=RO
         L=CO
         DO 340 I=1,K
340      ROM(I)=X(I,11)/CO
         PRINT 341, ROM
341      FORMAT(11H ROW MEANS /(1X,F15.3))
         DO 342 I=1,L
342      COM(I)=X(11,I)/RO
         PRINT 343, COM
343      FORMAT(13H COLUMN MEANS /(1X,F15.3/))
         DO 45 I=1,K
45       SSWC=SSWC + C(I,4)
         GPQ=X(11,11)**2/(RO*CO)
         DO 50 I=1,II
50       SSR=X(I,11)*X(I,11)+SSR
         SSR=SSR/CO
         DO 51 I=1,JJ
51       SSC=SSC+X(11,I)*X(11,I)
         SSC=SSC/RO
         DO 52 I=1,II
         DO 52 J=1,JJ
52       SSI=SSI+X(I,J)*X(I,J)
         PRINT 705, GPQ
         PRINT 706, SSWC
         PRINT 707, SSR
         PRINT 708, SSC
         PRINT 709, SSI
         BR=(SSR-GPQ)*HM
         BC=(SSC-GPQ)*HM
```

```
      AINT=(SSI-SSR-SSC+GPQ)*HM
      DO 55 I=1,K
      SD=SD+C(I,1)
      SSD=SSD+C(I,2)
55    EN=EN+C(I,3)
      TOT=SSD-(SD*SD)/EN
      DFR=CO-1.
          AMSR=BR/DFR
      DFC=RO-1.
      AMSC=BC/DFC
      AMSI=AINT/(DFR*DFC)
      DFI=DFR*DFC
      DO 60 I=1,K
60    DFE=DFE+C(I,3)-1.
      AMSE=SSWC/DFE
      PRINT 710
      PRINT 711, BR,DFR,AMSR
      PRINT 711, BC,DFC,AMSC
      PRINT 711, AINT,DFI,AMSI
      PRINT 711, SSWC,DFE, AMSE
      PRINT 712, TOT
701   FORMAT (20H SUMMING CHECK CELL  , I4, 5X, 4F20.3)
702   FORMAT(17H0HARMONIC MEAN = ,F10.5)
703   FORMAT(22H0MEANS AND SUMS MATRIX     )
704   FORMAT(10(5X,10F11.1,F13.1/)//5X,10F11.1,F13.1)
705   FORMAT(15H0COMP TERM 1 = ,F15.5)
706   FORMAT(15H0COMP TERM 2 = ,F15.5)
707   FORMAT(15H0COMP TERM 3 = ,F15.5)
708   FORMAT(15H0COMP TERM 4 = ,F15.5)
709   FORMAT(15H0COMP TERM 5 = ,F15.5)
710   FORMAT(1H0,9X,12H SUM SQUARES,17X,5H D.F.,17X,12H MEAN SQUARE)
711   FORMAT(13X,F12.4,17X,F5.0,17X,F12.4)
712   FORMAT(10X,F12.4)
      STOP
      END
      END
```

Program QGRPS

This program will give the Expert correlation and z' conversions among any pairs of Q sorts, and will give the "Adjustment" score (calculated on the basis of the Expert sort) for each Q sort.

Limits:

Q sorts are to be read in groups of not more than 50. As many groups may be read in as desired.

Pair correlations may be obtained only within groups. These pairs may be called for explicitly, or the program may be asked to compute the r.'s and z' conversions for each possible pair within a given group. No more than 600 pairs may be called explicitly, within any single group.

Card preparation:

1. First control card.
   This card contains the identification of the problem (e.g., the particular group of Q sorts being considered), the number of Q sorts in the problem, and instructions concerning Expert and Adjustment (Adj.) socres.

1)	Cols. 1-16	Identification. This may be either alphabetic or numeric.
2)	Cols. 17-20	Number of Q sorts in problem. This number must be punched as far to the right as possible.
3)	Col. 22	"Expert" instructions. If Expert r.'s are wanted, punch the digit '1' in this column.
4)	Col. 24	"Adj." instructions. If Adj. scores are wanted, punch the digit '1' in this column.

2. Pair control cards.
   These cards control the pairing of Q sorts to be correlated. If no explicit pairs are called for, punch '-1' in columns 1 and 2 of the pair control card. This will yield r.'s for all possible pair combinations within that group.

   If explicit pairs are called, they are to be indicated on a card or series of cards. A given Q sort is indicated by the numerical order in which it is read in as data, and is referred to on the pair control cards by this number. A set of two numbers thus indicates a pair of Q sorts to be correlated. No more than 15 pairs may be indicated on a given pair control card, and each pair control card must be filled with its fifteen pairs before a new pair control card is begun.

Proper spacing of the pair indications on the pair control cards is imperative. Each pair is allotted 5 columns. The first two are for the first member of the pair, the second two for the second member, and the fifth is a blank. Thus, to pair Q sorts #2 and #3, for instance, as the first correlation in the problem, the first 5 columns of the first pair control card would read: *2*3* (where '*' indicates a blank space).

Columns 1 and 2, and columns 3 and 4, respectively, of any set of 5 columns on the pair control cards, always indicate the 1st and 2nd members of a pair to be correlated, and in these 4 columns, a blank is read as zero. Therefore, in the above example, if you punched: 2*3**, the correlation given would be for #'s 20 and 30 rather than for #'s 2 and 3.

Pair control cards begin in the first column and end (with a blank space) in any column which is a multiple of 5, up to column 75. An example of a set of pair control cards might be:

```
*1*2**1*3**1*4**1*5**1*6**110**111**210**211**310**4*5**6
*7**8**920*020*1112*1314*1516*1718*1920*20*2*20*3*20*4*20*5*
```

These cards would yield the following pairs: 1 and 2, 1 and 3, 1 and 4, 1 and 5, 1 and 6, 1 and 10, 1 and 11, 2 and 10, 2 and 11, and so forth. No more than 600 pairs are allowed for a given problem.

3. Data cards.
Each Q sort requires 3 data cards. The first 8 columns of each of these cards are reserved for identification, consisting of alphabetic and/or numeric information which will be printed out with the answers. On the first card of each set of 3 are punched the first 35 items of the Q sort (from the triangular array data sheets), in columns 9-78. This means that there are only 2 spaces per item, and no blank spaces between items. A blank space is read as zero.

The second card of the set is identical to the first, except that it contains items 36-70, punched in columns 9-78.

The third card contains items 71-80, punched in columns 9-38.

Q sorts are placed in order, and this order determines the way in which they will be called in pairs for correlation (see section 2).

4. For a second (and further) problem all control cards must be repeated, in the same order as for the first problem.

5. A blank card is inserted following the last data card of the last group.

Print Out.

All correlations (r. scores) and their z' conversions are identified by the identification on the pair of Q sorts correlated. Thus if 2 Q sorts have in their first 16 columns JUDY*F** and TOM*S***, respectively, their correlation with each other will be indicated by:

```
JUDY*F** AND TOM*S** R = xxxxxxxx  ZCONV = xxxxxxxx
```

Correlation with Expert will be indicated in the same maqner, as will the Adj. score.

In addition, the pairs called for will be printed out under the heading "NPICK ARRAY." The Q sort scores converted to the score listing found on the Q sort score sheet will be printed out under the heading "SCORE ARRAY." The sum, sum of squares, sum times number of items (80), and sum squared, for each Q sort, will be printed out.

This data will be repeated for each problem.

<u>Note</u>: This program utilizes an extra card following the final END card. This card contains the "EXPERT" sort, and is treated by the program as a data card, it should be kept as a part of the program deck.

```
      PROGRAM QGRPS2
      DIMENSION NPICK(2,600), SUMS(4,51), SCORE(81,51), RAW(81)
      DIMENSION IDENT(2)
      CONTROL = 0.
      DO 7777 I=1,81
      DO 7777 J=1,51
7777   SCORE(I,J)=4.
      READ 105, (SCORE(I,51), I=2,81)
102   READ 100, IDENT, NGRP, EXPERT, ADJ
      IF (NGRP) 91, 91, 101
101   PRINT 83, IDENT
      IP = 1
      JP = 15
20    READ 21, ((NPICK(I,J), I=1,2), J=IP,JP)
      IF (NPICK(1,IP)) 29,22,23
23    IP = IP + 15
      JP = JP + 15
      GO TO 20
29    CONTROL = 1.
22    CONTINUE
      PRINT 82, ((NPICK(I,J), I=1,2), J=1,JP)
82    FORMAT (1H0,11HNPICK ARRAY/(10X,15(2I2,1X)))
11    DO 10 K = 1, NGRP
      READ 1, RAW
      SCORE (1,K) = RAW(1)
      DO 2 I = 2,4
      JJ = RAW(I)+1.
2     SCORE(JJ,K) = 0.
      DO 3 I = 5,10
      JJ = RAW(I)+1.
3     SCORE(JJ,K) = 1.
      DO 4 I = 11,19
      JJ = RAW(I)+1.
4     SCORE(JJ,K) = 2.
      DO 5 I = 20, 32
      JJ = RAW(I)+1.
5     SCORE(JJ,K) = 3.
      DO 6 I = 33,50
      JJ = RAW(I)+1.
6     SCORE(JJ,K)=4.
      DO 7 I = 51,63
      JJ = RAW(I)+1.
7     SCORE(JJ,K) = 5.
      DO 8 I = 64,72
      JJ = RAW(I)+1.
8     SCORE(JJ,K) = 6.
      DO 9 I = 73, 78
      JJ = RAW(I)+1.
9     SCORE(JJ,K) = 7.
      DO 10 I = 79,81
      JJ = RAW(I)+1.
10    SCORE(JJ,K) = 8.
```

```
      PRINT 81, ((SCORE(I,J), I=1,81), J=1,NGRP)
81    FORMAT (1H0,11HSCORE ARRAY/(10X,A8,1X,40F2.0/19X,40F2.0))
      DO 13 K = 1,NGRP
      SUMX = 0.
      SUMXSQ = 0.
      DO 12 I = 2,81
      SUMX = SUMX + SCORE(I,K)
12    SUMXSQ = SUMXSQ + SCORE(I,K)*SCORE(I,K)
      SUMS(1,K) = SUMX
      SUMS(2,K) = SUMXSQ
      SUMS (3,K) = 80.*SUMXSQ
13    SUMS(4,K) = SUMX*SUMX
      PRINT 84
84    FORMAT(74H0SUMS(1) IS SUMX, SUMS(2) IS SUMXSQ, SUMS(3) IS NSUMX, S
     1UMS(4) IS SUMXQSQ.)
      PRINT 85, ((SUMS(I,J), I=1,4), J=1,NGRP)
85    FORMAT (5X,4F12.3)
      IF (CONTROL) 45,45,35
45    DO 30 I = 1,JP
      SUMXY = 0.
      IF (NPICK(1,I)) 90, 90, 31
31    DO 32 JJ = 2,81
32    SUMXY = SUMXY + SCORE(JJ,NPICK(1,I))*SCORE(JJ,NPICK(2,I))
      XNUM = 80.*SUMXY-SUMS(1,NPICK(1,I))*SUMS(1,NPICK(2,I))
      COMPX = 80.*SUMS(2,NPICK(1,I))-SUMS(1,NPICK(1,I))*SUMS(1,NPICK(
     11,I))
      COMPY = 80.*SUMS(2,NPICK(2,I))-SUMS(1,NPICK(2,I))*SUMS(1,NPICK
     1(2,I))
      DENOM = SQRTF(COMPX*COMPY)
      R = XNUM/DENOM
      RRR=ABSF(R-1.)
      IF(RRR-.001)910,910,900
910   Z=9.
      GO TO 901
900   Z=.5*(LOGF(1.+ABSF(R))-LOGF(1.-ABSF(R)))
901   PRINT 200
200   FORMAT (1H0)
      M=NPICK(1,I)
      N=NPICK(2,I)
30    PRINT 87, SCORE(1,N), SCORE(1,M), R, Z
      GO TO 90
35    J = NGRP-1
      DO 41 K1 = 1,J
      L = K1+1
      DO 41 K2 = L,NGRP
      SUMXY = 0.
      DO 42 I = 2,81
42    SUMXY = SUMXY+SCORE(I,K1)*SCORE(I,K2)
      XNUM = 80.*SUMXY-SUMS(1,K1)*SUMS(1,K2)
      COMPX = 80.*SUMS(2,K1)-SUMS(1,K1)*SUMS(1,K1)
      COMPY = 80.*SUMS(2,K2)-SUMS(1,K2)*SUMS(1,K2)
      DENOM = SQRTF(COMPX*COMPY)
```

Note: Statement 84 needs "X, space S" in columns 69-72.

```
      R = XNUM/DENOM
      RRR=ABSF(R-1.)
      IF (RRR-.001)810,810,800
810   Z=9.
      GO TO 801
800   Z=.5*(LOGF(1.+R)-LOGF(1.-R))
801   PRINT 200
41    PRINT 87, SCORE(1,K1), SCORE(1,K2), R, Z
90    IF (ADJ) 400, 400, 300
300   DO 301 K=1,NGRP
      SADJ = 0.
      DO 304 I=2,81
      IF (SCORE(I,K)-4.) 303,304,305
303   IF (SCORE(I,51)-4.) 306, 304, 304
305   IF (SCORE(I,51)-4.) 304,304,306
306   SADJ = SADJ + 1.
304   CONTINUE
      PRINT 200
301   PRINT 307, SCORE (1,K), SADJ
307   FORMAT (1H ,A8,14H  ADJ SCORE =  ,F4.1)
400   IF(EXPERT) 40,40,70
70    SUMY = 0.
      SUMYSQ = 0.
      DO 71 I = 2,81
      SUMY = SUMY+SCORE(I,51)
71    SUMYSQ = SUMYSQ + SCORE(I,51)*SCORE(I,51)
      COMPY = 80.*SUMYSQ-SUMY*SUMY
      DO 79 J = 1,NGRP
      SUMXY = 0.
      DO 78 I = 2,81
78    SUMXY = SUMXY+SCORE(I,J)*SCORE(I,51)
      XNUM = 80.*SUMXY-SUMS(1,J)*SUMY
      COMPX = 80.*SUMS(2,J)-SUMS(1,J)*SUMS(1,J)
      DENOM = SQRTF(COMPX*COMPY)
      R = XNUM/DENOM
      Z=.5*(LOGF(1.+R)-LOGF(1.-R))
      PRINT 200
79    PRINT 88, SCORE (1,J), R, Z
40    GO TO 102
100   FORMAT (2A8,I4,2F2.0)
83    FORMAT (1H2,2A8)
21    FORMAT(15(2I2,1X))
1     FORMAT(A8,35F2.0/8X,35F2.0/8X,10F2.0)
87    FORMAT(1H ,A8,5H AND ,A8,5H R = ,F5.4,8H ZCONV = ,F6.4)
88    FORMAT (13H EXPERT WITH ,A8,F5.4,8H ZCONV = ,F6.4)
105   FORMAT(80F1.0)
91    STOP
      END
      END
```

Program TPAIRS

The t test is one of the most ubiquitous of statistical techniques in psychological and educational research. While a single t test is not particularly difficult to carry out by hand, especially with the aid of a desk calculator, the experimenter often wants t values for comparisons over a number of different variables. Moreover, he is often interested in splitting up his total experimental group in several different ways in order to test various hypotheses by means of t tests.

For example, a number of ability tests might have been administered to a class of 6th graders. The experimenter wants to compare boys and girls of each of these tests; he wnats to compare over- and under-achievers (previously defined), and he wants to compare children from high and from low socioeconomic level backgrounds. If five tests have been given, this means fifteen separate t values to be found.

Program TPAIRS is designed to handle t tests for up to 30 variables, and in varying combinations or groupings of up to 400 subjects.

In order to describe the way the program is used, certain terminology must be clarified. First, every t test between means may be thought of as a comparison between two treatment or classification groups: boys vs. girls, high achievers vs. low achievers, individual attention vs. group recitation, etc. These two sets of scores, corresponding to the two sets of dependent variable scores, will be referred to as the X and the Y variables. The letters "X" and "Y" have no particular significance; they simply make it possible to distinguish between the two groups which are being compared.

Secondly, as a total set of data is being considered, it can be seen that the various divisions to which it is subjected will divide it into smaller and smaller "basic groups." In our example of the 6th grade class, the boy/girl comparison creates two groups. The high-achiever/low-achiever comparison creates two different groups, yielding four "basic groups": boy high-achievers, boy low-achievers, girl high-achievers, and girl low-achievers. The socioeconomic level division gives 8 basic groups: boy high-achiever, high socioeconomic; boy high achiever, low socioeconomic; boy low-achiever, high socioeconomic; and so on. In other words, the basic group is the smallest grouping possible when all of the compari sons are combined--or the largest group that does not have to be re-divided.

Program TPAIRS is based on the notion of combining and re-combinig basic groups. The entire set of data is divided into its basic groups, and control cards indicate the combinations in which the experimenter is interested. Going back to our now-familiar example, a diagram of the class might look like this:

boys	high-achiever	high	socioeconomic	level	group 1-10 students
		low	"	"	group 2-12 students
	low-achiever	high	"	"	group 3-14 students
		low	"	"	group 4- 9 students
girls	high achiever	high	"	"	group 5-15 students
		low	"	"	group 6- 8 students
	low achiever	high	"	"	group 7-20 students
		low	"	"	group 8-17 students

Obviously, the experimenter is not interested in all possible combinations of these 8 groups. In fact, his interest lies in in just three combinations: 1, 2, 3, and 4 vs. 5, 6, 7 and 8; 1, 3, 5, and 7 vs. 2, 4, 6, and 8; and 1, 2, 5 and 6 vs. 3, 4, 7, and 8. He can so specify by means of control cards preceding his data deck. Moreover, he could, if we wished, test any subgrouping in which he was interested; high-achiever boys vs. low-achiever boys (1 and 2 vs. 3 and 4), for instance, or high socioeconomic girls vs. high socioeconomic boys (1 and 3 vs. 5 and 7).

With these terms defined, we can now proceed to the directions for using program TPAIRS.

Limits:

1. The total number of subjects may not exceed 400.
2. The total number of basic groups may not exceed 98.
3. No more than 10 basic groups for each X or Y variable set.
4. The total number of variables may not exceed 30.

Control cards and data:

1. Identification card. Following the program deck, the first card to be used is an identification card. This is simply printed out with the results of the t tests, and is not used in any of the actual computation. Any characters may be punched anywhere on the card, and will be printed out exactly as punched.

2. First control card.
   columns 1-4: number of variables to be used, punched as far to the right as possible.
   columns 5-8: total number of subjects (or, sum of the numbers in each basic group), punched as far to the right as possible.

3. Group control cards. These cards indicate the basic groups to be used for each t comparison. The groups composing the X variable for each test are given first, followed by the groups composing the Y variable.

    1) X group control card(s). In fields of 2 columns each, starting with columns 1 and 2, are the numbers of the basic groups composing the X variable for the first t test. Two blank columns indicate the end of one t test. All X variable sets must be listed together. Since 2 columns are used for each basic group or space indicator, there is room for 40 pieces of information on one card. Should this not be enough, another card or cards may be used. There is no limit on the number of group control cards, provided the limit of 25 separate groupings for t tests, with no more than 10 basic groups in any one X or Y variable combination, is not exceeded. Also, all of the basic groups within the X variable listings (and within the Y variable listings) must be consecutive; e.g., there must be no blank columns other than those indicating the beginning of a new combination, and those single blanks immediately preceding a single-digit number.

        Immediately following the last basic group of the last X variable combination, 99. There must be no blank columns between the number designating the final group and the 99.

    2) Separator card. After the last X group control card, a card with -1 in columns 1 and 2.

    3) Y group control card(s). The same as the X group control cards, but now listing the basic groups composing the Y variable part of the t test.

    4) Separator card. After the last Y group control card, a card with -1 in columns 1 and 2.

4. Format. The form of the data is described by an ordinary FORTRAN format statement, omitting the word "FORMAT," and with the first and last parentheses anywhere on the card.

5. Data cards. Data follows the pattern described by the format cards. Data must be in order

    $$X_{i_1 j_1 k_1}, X_{i_2 j_1 k_1}, X_{i_3 j_1 k_1}, \ldots X_{i_n j_1 k_1}, X_{i_1 j_2 k_1}, \ldots X_{i_n j_n k_n}$$

    where i designates the variables which are to be tested and which are referred to on the first control card. In other words, all the variables pertaining to subject 1 must be read in before any variables pertaining to subject 2 are read in.

Data must be ordered according to basic groups. Between each basic group, a blank card (or cards, if more than one card is called for in the FORMAT statement) must be placed.

Negative numbers are allowed in the data; however, no subject's scores may sum to exactly zero ($\Sigma X \neq 0$, where X refers to variables used in the t tests).

Output. Output for TPAIRS is somewhat redundant, but this redundancy does provide a means of checking out the rather complicated input.

1) Arrays NXS and NYS. Here are printed, in columnar form, the groupings set up on the group control cards.

2) DATA array. An echo check of the data itself, with basic groups now labeled with their respective group numbers.

3) Identification of problem, as punched on identification card.

4) For each variable, and for each pair of X-Y combinations listed, the values of t, mean of X, mean of Y, n of x, n or Y, and σ_{diff}.

```
      PROGRAM TPAIRS
      DIMENSION TRAK(60,3,30),FMT(10),IDENT(10)
      DIMENSION DATA(400,31),ONTO(320),NXS(11,26),NYS(11,26)
      COMMON DATA
      READ 5, IDENT
5     FORMAT (10A8)
      PRINT 6,IDENT
6     FORMAT (1H ,10A8)
      READ 7, NVZ,NSS
7     FORMAT (2I4)
      M=1
      N=40
      JUMP=0
9     READ 10,(ONTO(I),I=M,N)
10    FORMAT (40F2.0)
      DO 11 L=M,N
      IF(ONTO(M))118,19,19
19    IF(ONTO(L)-99.) 11,18,18
11    CONTINUE
18    M=M+40
      N=N+40
      GO TO 9
118   K=1
      I=1
      J=1
13    IF (ONTO(K)) 17,17,14
14    IF (ONTO(K)-99.) 16,15,15
15    GO TO 20
16    IF (JUMP) 23,23,21
21    NYS(I,J)=ONTO(K)
      GO TO 22
23    NXS(I,J)=ONTO(K)
22    K=K+1
      I=I+1
      GO TO 13
17    I=1
      J=J+1
      K=K+1
      GO TO 13
20    IF (JUMP) 24,24,25
24    M=1
      N=40
      JUMP = 1
      GO TO 9
25    CONTINUE
      PRINT 149,((NXS(I,J),J=1,25),I=1,10)
149   FORMAT(10H0ARRAY NXS/(10X,25I3))
      PRINT 150, ((NYS(I,J),J=1,25),I=1,10)
150   FORMAT(10H0ARRAY NYS/(10X,25I3))
      READ 90,FMT
90    FORMAT (10A8)
      L=1
      K=1
      NV=NVZ+1
3     DATA(K,1)=L
      READ FMT,(DATA(K,J),J=2,NV)
      SUM=0.
      DO 70 J=2,NV
70    SUM=SUM+ABSF(DATA(K,J))
      IF (SUM) 4,4,28
28    IF (K-NSS) 2,29,29
2     K=K+1
      GO TO 3
4     L=L+1
      GO TO 3
29    CONTINUE
      PRINT 151, ((DATA(I,J),J=1,31),I=1,NSS)
151   FORMAT(11HODATA ARRAY/(10X,F3.1,15F5.2/13X,15F5.2))
      DO 104 J=2,NV
      N=J-1
      I=1
      K=1
      A=1.
135   SUMX=0.
      SUMXSQ=0.
      NX=0
131   IF(DATA(I,1)-A) 102,100,101
100   SUMX=SUMX+DATA(I,J)
      SUMXSQ=SUMXSQ+DATA(I,J)*DATA(I,J)
      NX=NX+1
      I=I+1
      IF (NSS-I) 101,131,131
101   TRAK(K,1,N)=SUMX
      TRAK(K,2,N)=SUMXSQ
      TRAK(K,3,N)=NX
      IF (NSS-I) 104,2007,2007
2007  K=K+1
      A=A+1.
      GO TO 135
102   PRINT 115
115   FORMAT(25H ERROR IN STORE SUMS LOOP)
      STOP 99
104   CONTINUE
      PRINT 153,(((TRAK(I,J,K),J=1,3),I=1,10),K=1,10)
153   FORMAT(11HOTRAK ARRAY/(10X,3F15.3))
      DO 60 J=2,NV
      JA=1
      JB=1
      L=0
50    SUMX=0.
      SUMXSQ=0.
      ANX=0.
      SUMY=0.
      SUMYSQ=0.
      ANY=0.
      IA=0
      IB=0
```

```
40     IA=IA+1
       IF(NXS(IA,JA)) 42,42,41
41     M=NXS(IA,JA)
       N=J-1
       SUMX=SUMX+TRAK(M,1,N)
       SUMXSQ=SUMXSQ+TRAK(M,2,N)
       ANX=ANX+TRAK(M,3,N)
       GO TO 40
42     IB=IB+1
43     IF(NYS(IB,JB)) 45,45,44
44     M=NYS(IB,JB)
       SUMY=SUMY+TRAK(M,1,N)
       SUMYSQ=SUMYSQ+TRAK(M,2,N)
       ANY=ANY+TRAK(M,3,N)
       GO TO 42
45     AVGX=SUMX/ANX
       AVGY=SUMY/ANY
       COMPX=SUMXSQ-SUMX*SUMX/ANX
       COMPY=SUMYSQ-SUMY*SUMY/ANY
       COMP=(COMPX+COMPY)/(ANX+ANY-2.)*((1./ANX;+(1./ANY))
       SIGMA=SQRTF(COMP)
       T=ABSF(AVGX-AVGY)/SIGMA
       L=L+1
       PRINT 110, L, N, T
110    FORMAT(10HOT OF SET ,I3,11H, VARIABLE ,I3,4H, = ,F6.3)
       PRINT 111, AVGX, AVGY
111    FORMAT(10X,12HMEAN OF X = ,F10.3/10X,12HMEAN OF Y = ,F10.3)
       PRINT 171, ANX, ANY
171    FORMAT (10X,9HN OF X = ,F5.0/10X,9HN OF Y = ,F5.0)
       PRINT 112, SIGMA
112    FORMAT (10X,8HSIGMA = ,F10.3)
       JA=JA+1
       JB=JB+1
       IF(NXS(1,JA)) 60,60,61
61     GO TO 50
60     PRINT 113
113    FORMAT (1H2)
       STOP
       END
       END
```